

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
29. Juli 2004 (29.07.2004)

PCT

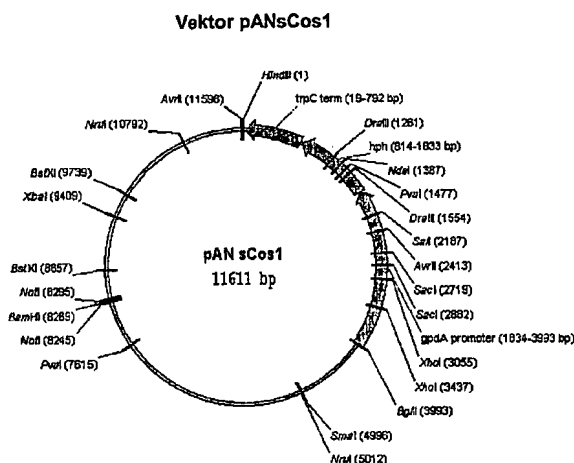
(10) Internationale Veröffentlichungsnummer
WO 2004/063359 A2

- (51) Internationale Patentklassifikation⁷: C12N 1/15, 15/80, C12P 23/00, A23J 1/00, 3/00, A23L 1/28, 1/275
- (21) Internationales Aktenzeichen: PCT/EP2004/000099
- (22) Internationales Anmeldedatum:
9. Januar 2004 (09.01.2004)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:
103 00 649.4 9. Januar 2003 (09.01.2003) DE
103 41 271.9 8. September 2003 (08.09.2003) DE
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF AKTIENGESELLSCHAFT [DE/DE]; 67056 Ludwigshafen (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): MATUSCHEK, Markus [DE/DE]; Karolinenstr. 5, 69469 Weinheim (DE). KLEIN, Daniela [DE/DE]; M 7, 2, 68161 Mannheim (DE). HEINEKAMP, Thorsten [DE/DE]; Alte Ziegelei 1B, 30419 Hannover (DE). SCHMIDT, Andre [DE/DE]; Magdeburger Str. 11, 31832 Springe (DE). BRAKHAGE, Axel [DE/DE]; Schneiderberg 58, 30167 Hannover (DE). ACHATZ, Brigitte [DE/DE]; Windeckstr. 26, 68163 Mannheim (DE).
- (74) Anwalt: FITZNER, Uwe; Lintorfer Str. 10, 40878 Ratingen (DE).
- (81) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES,

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR PRODUCING CAROTENOIDS OR THEIR PRECURSORS USING GENETICALLY MODIFIED ORGANISMS OF THE *BLAKESLEA* GENUS, CAROTENOIDS OR THEIR PRECURSORS PRODUCED BY SAID METHOD AND USE THEREOF

(54) Bezeichnung: VERFAHREN ZUR HERSTELLUNG VON CAROTINOIDEN ODER DEREN VORSTUFEN MITTELS GENTECHNISCH VERÄNDERTER ORGANISMEN DER GATTUNG *BLAKESLEA*, MIT DEM VERFAHREN HERGESTELLTE CAROTINOIDE ODER DEREN VORSTUFEN UND DEREN VERWENDUNG



(57) Abstract: The invention relates to a method for producing carotenoids or their precursors using genetically modified organisms of the *Blakeslea* genus. Said method comprises the following steps (i) transformation of at least one of the cells, (ii) optional homokaryotic conversion of the cells obtained in step (i) to produce cells, in which one or more genetic characteristics of the nuclei are all modified in an identical manner and said modification manifests itself in the cells, (iii) selection and reproduction of the genetically modified cell or cells, (iv) cultivation of the genetically modified cells, (v) preparation of the carotenoids produced by the genetically modified cells or the carotenoid precursor produced by said genetically modified cells. The invention also relates to carotenoids or their precursors produced according to said method and to the use thereof.

[Fortsetzung auf der nächsten Seite]

WO 2004/063359 A2



FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK,

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(57) Zusammenfassung: Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderter Organismen der Gattung *Blakeslea* umfassend (i) (i) Transformation mindestens einer der Zellen, (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder in mehreren genetischen Merkmalen alle gleichartig verändert sind und diese genetische Veränderung zur Ausprägung bringen, und (ix) Selektion und Vermehrung der gentechnisch veränderten Zelle oder Zellen, (x) Kultivierung der gentechnisch veränderten Zellen, (xi) Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe; nach dem Verfahren hergestellte Carotinoide oder deren Vorstufen und deren Verwendung.

**Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen
mittels gentechnisch veränderter Organismen der Gattung Blakes-
lea, mit dem Verfahren hergestellte Carotinoide oder deren Vorstufen
und deren Verwendung**

5

Die Erfindung betrifft ein Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderten Organismen der Gattung Blakeslea, mit dem Verfahren hergestellte Carotinoide oder deren Vorstufen und deren Verwendung und Bereitstellung, besonders als hoch-
10 reine Carotinoide, als Nahrungsmittel, enthaltend Carotinoideproduzierende Organismen und mindestens ein Carotinoid, insbesondere Tierfuttermittel, Tierfutterergänzungsmittel und Nahrungsergänzungsmittel, sowie die Verwendung der aus dem Verfahren erhältlichen Carotinoide zur
15 bereitungen, Nahrungsmitteln oder Nahrungsergänzungsmitteln.

Blakeslea trispora ist als Produktionsorganismus für β -Carotin (Ciegler, 1965, Adv Appl Microbiol. 7:1) und Lycopin bekannt (EP 1201762, EP 1184464, WO 03/038064).

20

Von Blakeslea trispora sind bisher verschiedene DNA-Sequenzen bekannt, insbesondere die DNA-Sequenz, die für die Gene der Carotinoidbiosynthese von Geranylgeranylpyrophosphat bis β -Carotin codiert (WO 03/027293).

25

Insbesondere aufgrund der hohen Produktivität, die mit Blakeslea in der Produktion von Lycopin und β -Carotin erreicht werden, bietet sich dieser Organismus zur fermentativen Herstellung von Carotinoiden an.

30

Es ist auch von Interesse die Produktivitäten der bisher natürlicherweise produzierten Carotine und deren Vorstufen weiter zu steigern und die Her-

stellung weiterer Carotinoide, wie z. B. Xanthophylle zu ermöglichen, die von Blakeslea bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

- 5 Carotinoide werden Futtermitteln, Nahrungsmitteln, Nahrungsergänzungsmitteln, Kosmetika und Arzneimitteln zugesetzt. Die Carotinoide dienen vor allem als Pigmente zur Färbung. Daneben werden die antioxidative Wirkung der Carotinoide und andere Eigenschaften dieser Substanzen genutzt. Man unterteilt die Carotinoide in die reinen Kohlenwasserstoffe, die Carotine und die sauerstoffhaltigen Kohlenwasserstoffe, die Xanthophylle. Xanthophylle wie Canthaxanthin und Astaxanthin werden beispielsweise zur Pigmentierung von Hühnereiern und Fischen eingesetzt (Britton et al. 1998, Carotinoids, Vol 3, Biosynthesis and Metabolism). Die Carotine β -Carotin und Lycopin werden vor allem in der Humanernährung eingesetzt. β -Carotin wird beispielsweise als Getränkefarbstoff verwendet. Lycopin hat eine krankheitsvorbeugende Wirkung (Argwal und Rao, 2000, CMAJ 163:739-744; Rao und Argwal 1999, Nutrition Research 19:305-323). Die farblose Carotinoidvorstufe Phytoen kommt vor allem für Anwendungen als Antioxidans in kosmetischen, pharmazeutischen oder dermatologischen Zubereitungen in Frage.

- Der überwiegende Teil der Carotinoide und deren Vorstufen, die als Zusatzstoffe für die oben genannten Anwendungen eingesetzt werden, wird durch chemische Synthese hergestellt. Die chemische Synthese ist technisch sehr aufwendig und verursacht hohe Herstellkosten. Fermentative Verfahren sind demgegenüber technisch verhältnismäßig einfach und basieren auf kostengünstigen Einsatzstoffen. Fermentative Verfahren zur Herstellung von Carotinoiden und deren Vorstufen können dann wirtschaftlich attraktiv und wettbewerbsfähig zur chemischen Synthese sein, wenn die Produktivität der bisherigen fermentativen Verfahren gesteigert würde oder neue Carotinoide auf Basis der bekannten

würde oder neue Carotinoide auf Basis der bekannten Produktionsorganismen hergestellt werden könnten.

5 Hierzu ist eine gentechnische, d. h. gezielte genetische Veränderung von Blakeslea erforderlich. Insbesondere, wenn Xanthophylle produziert werden sollen, da diese Verbindungen natürlicherweise vom Wildtyp der Blakeslea nicht synthetisiert werden.

10 Z. B. zur Herstellung von Phytoen mittels Fermentation von Blakeslea trispora sind bisher zwei Methoden bekannt:

(i) Durch zufallsabhängige Mutagenese mit chemischen Agenzien wie MNNG können Mutanten erzeugt werden, in denen Phytoen nicht zu Lycopin und somit nicht weiter zu β -Carotin umgesetzt werden kann (Mehta und Cerdá-Olmedo, 1995, Appl. Microbiol. Biotechnol. 42:836-838).

20 (ii) Durch Zugabe von Inhibitoren des Enzyms Phytoendesaturase wie z.B. Diphenylamin und Zimtalkohol kann die weitere Umsetzung von Phytoen blockiert werden, so dass es sich anreichert (Cerdá-Olmedo, 1989, In: E. Vandamme, ed. Biotechnology of vitamin, growth factor and pigment production. London: Elsevier Applied Science, S. 27-42).

25 Die genannten Methoden zur Herstellung von Phytoen mit Blakeslea trispora weisen jedoch eine Reihe von Nachteilen auf.

30 Die zufallsabhängige Mutagenese betrifft in der Regel nicht nur die Gene der Carotinoidbiosynthese zur weiteren Umsetzung von Phytoen, sondern auch weitere wichtige Gene. Daher sind Wachstum und Syntheseleistung der Mutanten oft beeinträchtigt. Die Erzeugung z. B. von Phytoenüberpro-

duzenten durch zufallsabhängige Mutagenese von Lycopinüberproduzenten oder β -Carotinüberproduzenten ist daher entweder nicht oder nur mit großem experimentellem Aufwand zu erreichen. Die Zugabe von Inhibitoren verursacht eine Erhöhung der Produktionskosten und gegebenenfalls
5 eine Verunreinigung des Produktes. Daneben kann das Zellwachstum durch den Inhibitor beeinträchtigt werden, so dass die Produktion von Carotinoiden oder deren Vorstufen, insbesondere Phytoen eingeschränkt wird.

10 Durch eine gentechnische Veränderung könnten die oben genannten Nachteile der zufallsabhängigen Mutagenese und der Inhibitorzugabe vermieden werden.

15 Allerdings sind bisher keine Methoden zur gentechnischen, d. h. gezielten gentechnischen Veränderung von *Blakeslea*, insbesondere *Blakeslea trispora* bekannt.

Als Methode zur Herstellung von gentechnisch veränderten Pilzen wurde
20 in einigen Fällen die Agrobacterium-vermittelte Transformation erfolgreich eingesetzt. So sind z. B. folgende Organismen durch Agrobakterien transformiert worden: *Saccharomyces cerevisiae* (Bundock et al., 1995, EMBO Journal, 14:3206–3214), *Aspergillus awamori*, *Aspergillus nidulans*, *Aspergillus niger*, *Colletotrichum gloeosporioides*, *Fusarium solani pisi*,
25 *Neurospora crassa*, *Trichoderma reesei*, *Pleurotus ostreatus*, *Fusarium graminearum* (van der Toorren et al., 1997, EP 870835), *Agraricus bisporus*, *Fusarium venenatum* (de Groot et al., 1998, Nature Biotechnol. 16:839–842), *Mycosphaerella graminicola* (Zwiers et al. 2001, Curr. Genet. 39:388–393), *Glarea lozoyensis* (Zhang et al., 2003, Mol. Gen. Ge-

nomics 268:645–655), *Mucor miehei* (Monfort et al. 2003, FEMS Microbiology Lett. 244:101 – 106).

5 Von Interesse ist besonders eine homologe Rekombination, bei der zwischen der einzuführenden DNA und der Zell-DNA möglichst viele Sequenzhomologien bestehen, so dass eine ortsspezifische Einführung bzw. Ausschaltung von genetischer Information im Genom des Empfängerorganismus möglich ist. Andernfalls wird die Spender-DNA durch illegitime bzw. nicht-homologe Rekombination ins Genom des Empfängerorganismus integriert, was nicht ortsspezifisch erfolgt.

Eine durch *Agrobacterium* vermittelte Transformation und anschließende homologe Rekombination der transferierten DNA wurde bisher bei folgenden Organismen nachgewiesen: *Aspergillus awamori* (Gouka et al. 1999, Nature Biotech 17:598-601), *Glarea lozoyensis* (Zhang et al., 2003, Mol. Gen. Genomics 268:645-655), *Mycosphaerella graminicola* ((Zwiers et al. 2001, Curr. Genet. 39:388-393).

20 Als weitere Methode zur Transformation von Pilzen ist die Elektroporation bekannt. Die integrative Transformation von Hefe durch Elektroporation wurde von Hill, Nucl. Acids. Res. 17:8011 gezeigt. Für filamentöse Pilze wurde die Transformation durch Chakabarty und Kapoor beschrieben (1990, Nucl. Acids. Res. 18:6737).

25 Eine „biolistische“ Methode, d.h. die Übertragung von DNA durch Beschuss von Zellen mit DNA-beladenen Partikeln wurde beispielsweise für *Trichoderma harzianum* und *Gliocladium virens* beschrieben (Lorito et al. 1993, Curr. Genet. 24:349–356).

Diese Methoden konnten bisher jedoch nicht erfolgreich zur gezielten genetischen Veränderung von *Blakeslea* und insbesondere *Blakeslea trispora* eingesetzt werden.

- 5 Eine besondere Schwierigkeit bei der Herstellung von gentechnisch veränderten *Blakeslea* und *Blakeslea trispora*, ist die Tatsache, dass deren Zellen in allen Stadien des sexuellen und des vegetativen Zellzyklus mehrkernig sind. In Sporen von *Blakeslea trispora* Stamm NRRL2456 und NRRL2457 wurden z. B. im Durchschnitt 4,5 Kerne pro Spore nachgewiesen (Metha und Cerdá-Olmedo, 1995, Appl. Microbiol. Biotechnol. 42:836–
10 838). Dies hat zur Folge, dass die gentechnische Veränderung in aller Regel nur in einem oder wenigen Kernen vorliegt, die Zellen also heterokaryotisch sind.
- 15 Sollen die genetisch veränderten *Blakeslea*, insbesondere *Blakeslea trispora* zur Produktion eingesetzt werden, so ist es insbesondere bei einer Gendelektion wichtig, dass in den Produktionsstämmen die gentechnische Veränderung in allen Kernen vorliegt, so dass eine stabile und hohe Syntheseleistung ohne Nebenprodukte möglich wird. Die Stämme müssen
20 folglich in Bezug auf die gentechnische Veränderung homokaryotisch sein.

Lediglich für *Phycomyces blakesleeae* ist ein Verfahren beschrieben worden, um homokaryotische Zellen zu erzeugen (Roncero et al., 1984, Mutat. Res. 125:195). Durch Zugabe des mutagenen Agens MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) werden nach dem dort beschriebenen
25 Verfahren Kerne in den Zellen eliminiert, so dass statistisch eine gewisse Anzahl von Zellen mit nur noch einem funktionellem Kern vorliegt. Die Zellen werden dann einer Selektion unterzogen, in der nur einkernige Zellen mit einem rezessiven Selektionsmarker zu einem Mycel auswachsen können. Die Nachkommen dieser selektierten Zellen sind mehrkernig und
30 homokaryotisch. Ein rezessiver Selektionsmarker für *Phycomyces blakes-*

leanus ist z. B. dar. dar⁺-Stämme nehmen das toxische Riboflavin-Analog 5-Carbon-5-deazariboflavin auf; dar⁻-Stämme dagegen nicht (Delbrück et al. 1979, Genetics 92:27). Die Selektion von rezessiven Mutanten erfolgt durch Zugabe von 5-Carbon-5-deazariboflavin (DARF).

5

Allerdings ist dieses Verfahren nicht für *Blakeslea*, insbesondere *Blakeslea trispora* bekannt und insbesondere nicht mit im Zusammenhang mit einer Transformation oder der Produktion von Carotinoiden oder deren Vorstufen beschrieben worden.

10

Auch die Isolierung aus natürlichen Quellen wird durchgeführt. Beispielsweise ist es für die Gewinnung von Phytoen bekannt, ein Gemisch aus Carotinoiden, Vitamin E und anderen Komponenten, welches auch Phytoen enthält, aus Tomaten, Karotten oder Palmöl usw. zu extrahieren.

15

Problematisch ist hierbei die Trennung der einzelnen Carotinoide voneinander. So ist beispielsweise das Phytoen nach diesem Verfahren nicht in reiner Form erhältlich. Insbesondere ist die natürlich vorkommende Menge der Carotinoiden in den Pflanzen gering.

20

Fermentative Verfahren sind demgegenüber technisch verhältnismäßig einfach und basieren auf kostengünstigen Einsatzstoffen. Fermentative Verfahren zur Herstellung von Carotinoiden können dann wirtschaftlich attraktiv und wettbewerbsfähig zur chemischen Synthese sein, wenn die Produktivität der bisherigen fermentativen Verfahren gesteigert würde oder

25

neue Carotinoide auf Basis der bekannten Produktionsorganismen hergestellt werden könnten. Problematisch bei der fermentativen Herstellung von Carotinoiden sind allerdings die Aufarbeitungsverfahren, die nur geringe Mengen an hochreinen Carotinoiden bereitstellen. Zudem sind dafür meist aufwendige Vielschritt-Prozesse ggf. unter Verwendung großer Lö-

30

sungsmittelmengen erforderlich. So fallen große Mengen Abfall an oder es

muss ein hoher Aufwand zur Wiederverwertung (Recycling) betrieben werden.

Die Produktion von Carotinoiden durch verschiedene Mikroorganismen ist an sich bekannt. So ist z. B. in der WO 00/13654 A2 offenbart, ein Gemisch aus Phytoen und Phytofluoren aus Algen der Art *Dunaliella* sp. zu extrahieren. Auch nach diesem Verfahren ist das Phytoen nicht in reiner Form erhältlich und muss von den anderen Produkten getrennt werden. Zudem handelt es sich um gentechnisch unveränderte Algen, deren Biosynthese mittels eines hinzugefügten Inhibitors beeinflusst werden muss.

Blakeslea trispora als Produktionsorganismus für β -Carotin ist auch aus der WO 98/03480 A1. bekannt. Hier werden β -Carotin Kristalle aus Biomasse von *Blakeslea trispora* mittels Extraktion erhalten. Allerdings müssen in dem beschriebenen Verfahren große Mengen unterschiedlicher Lösungsmittel eingesetzt werden, um Kristalle mit hoher Reinheit durch mehrere Extraktions- und Waschschriffe zu erhalten. Auch sind die erhaltenen Mengen β -Carotin bezogen auf die eingesetzte Menge Biomasse klein.

Aus der WO 01/83437 A1 ist ein Verfahren zur Extraktion von Astaxanthin aus Hefe bekannt, bei dem die Kulturbrühe zur Sterilisation und zum Zellaufschluss mit Mikrowellenstrahlung behandelt wird. Der Zellaufschluss mittels Mikrowellenstrahlung ist danach nötig, um Astaxanthin aus Hefe zu gewinnen, ohne es dabei zu zerstören. Anschließend soll Astaxanthin mittels Methanol, Ethanol oder Aceton oder deren Mischungen extrahiert werden. Hierzu sind allerdings große Mengen Lösungsmittel (5 bis 20 Teile Lösungsmittel auf 1 Teil Suspension) und ein langer Zeitraum (24h) erforderlich. Zudem sind keine Reinheiten des Astaxanthins angegeben und die erhaltenen Mengen sind klein. Versuche der Anmelderin und andere Veröffentlichungen bestätigen jedoch, dass eine Extraktion mittels Methanol oder Ethanol nicht durchführbar ist.

Aus der WO 98/50574 ist ebenfalls die Isolierung von Carotinoid Kristallen aus Biomasse von Mikroorganismen bekannt, wobei hiernach im Gegensatz zur WO 01/83437 A1 Methanol, Ethanol, Aceton nur zum Entfernen
5 von Lipiden aus der Biomasse d. h. .zum Waschen verwendet werden kann. Als Lösungsmittel zur Extraktion von Carotinoiden wird demnach Ethylacetat, Hexan oder ein Öl verwendet. Anschließend sind mehrere Reinigungs- und Waschschriffe mit großen Mengen Ethanol und Wasser nötig, wobei lediglich eine Reinheit von 93,3 % bei einer Ausbeute von 35
10 % erreicht wird.

Die WO 03/038064 A2 beschreibt die fermentative Produktion von Lycopin durch Co-Kultivierung von mutiertem *Blakeslea trispora* Paarungstyp (-) und *Blakeslea trispora* Paarungstyp (+), die ohne Zusatz von Inhibitoren
15 der Carotinoid Biosynthese Lycopin herstellen. Die Erzeugung der zur Fermentation eingesetzten Mutante wird durch unselektive chemische Mutation und anschließendes Screening vorgenommen. Die Aufarbeitung der Kulturbrühe erfolgt mittels Zellaufschluss und anschließender Reinigung mit unterschiedlichen wässrigen Medien mit verschiedenem Salzgehalt
20 und pH-Wert und mit Wasser nicht mischbaren organischen Lösungsmitteln wie Ethylacetat, Hexan und 1- Butanol zur Entfernung von Lipiden. Alternativ ist eine Extraktion mittels großer Mengen Ethylacetat beschrieben. Angaben zur Reinheit fehlen. Da Ethylacetat und Hexan Lösungsmittel für Lycopin sind, ist davon auszugehen, dass ein Teil des Lycopins
25 herausgewaschen und so die theoretische mögliche Ausbeute verringert wird.

Auch aus der WO 01/55100 A1 ist die Isolierung von Carotinoiden allgemein bzw. β -Carotin im speziellen aus der Biomasse durch Anwendung
30 mehrerer Wasch- und Reinigungsschritte auf die aufgeschlossene Biomasse ohne Extraktion mittels Lösungsmittel beschrieben. Hierzu wird

aufgeschlossene Biomasse von *Blakeslea trispora* mit Wasser, Lauge, Säure, Butanol und Ethanol gewaschen, so daß eine große Zahl unterschiedlicher Lösungsmittel und wässriger Medien verwendet werden muss. Die Reinheit des erhaltenen β -Carotins beträgt 96 – 98 %. Angaben zur Ausbeute fehlen jedoch.

Die WO 97/36996 A2 beschreibt allgemein ein Verfahren zur Isolierung von Substanzen (u. a. Carotinoide) aus Mikroorganismen, wobei die Substanzen aus der Biomasse mittels Fest/Flüssig-Extraktion isoliert werden. Ein Zellaufschluß soll hierbei nicht nötig sein, jedoch muss die Biomasse zunächst durch Extrusion in eine granulierten, poröse Gestalt gebracht werden. Wie nur Carotinoide isoliert werden können und wie deren Reinheit bzw. Ausbeute ist, ist nicht angegeben. Der Rückstand der Extrusion kann anschließend als Futtermittelzusatz verwendet werden.

In allen oben beschriebenen Verfahren müssen große Mengen Lösungsmittel zur Extraktion eingesetzt werden, um die isolierte Menge an Carotinoid durch vollständige Extraktion zu erhöhen, und/oder große Mengen wässriger Medien zur Reinigung und zum Waschen eingesetzt werden. Dies bedingt hohe Kosten und aufwendige Maßnahmen zur Wiederverwendung bzw. ggf. Abfälle.

Zudem werden die nahrhafte Kulturbrühe und die darin enthaltene Biomasse nach Extraktion bzw. Isolierung der Carotinoide als Abfall behandelt. Die oben angegebenen Verfahren haben neben diesen vordergründigen Nachteilen einen entscheidenden weiteren Nachteil. Es ist nämlich danach notwendig, die Carotinoide den Nahrungsmitteln nachträglich zuzusetzen, d. h. sie sind nicht Bestandteil der Nahrungsmittel an sich bzw. nicht in ausreichender Menge. Von großem Vorteil wäre daher, wenn der Gehalt an Carotinoiden in den Nahrungsmitteln bereits durch die eigentlichen Nahrungsmittel selbst gedeckt würde.

Es ist ebenfalls nötig die Produktivitäten der bisher natürlicherweise produzierten Carotine und deren Vorstufen weiter zu steigern und die Herstellung weiterer Carotinoide, wie z. B. Xanthophylle besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin zu ermöglichen, die von den Wildtypen der Mikroorganismen bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

Aufgabe der Erfindung ist es gentechnisch veränderte Zellen von *Blakeslea*-Stämmen, insbesondere *Blakeslea trispora* bereitzustellen, die Carotinoide oder deren Vorstufen, insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin produzieren. Zudem soll das Verfahren die Steigerung der Carotinoid-Produktivität der veränderten Zellen gegenüber den korrespondierenden Wildtypen erlauben. Ferner soll das Verfahren die Erzeugung neuer Zellen oder aus ihnen bestehendes Mycel erlauben, die sich für die Verwendung zur Herstellung von Carotinoiden oder deren Vorstufen eignen, die bisher nicht in wirtschaftlich interessanten Mengen aus den natürlich vorkommenden Pilzen gewinnbar waren, insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin. Das Verfahren soll dabei eine gentechnische Veränderung von *Blakeslea*-Stämmen, insbesondere *Blakeslea trispora* möglich machen und die Herstellung homokaryotischer gentechnisch veränderter Produktions-Stämme erlauben.

Des weiteren soll das Verfahren die Herstellung weiterer Carotinoide, wie z. B. Xanthophylle, insbesondere Astaxanthin oder Zeaxanthin und Phytoen oder Bixin ermöglichen, die von den Wildtypen der Mikroorganismen bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

Ferner ist es Aufgabe der vorliegenden Erfindung ein Verfahren zur Herstellung von Carotinoiden aus gentechnisch veränderte Zellen von Blakeslea-Stämmen, insbesondere *Blakeslea trispora*, zur Verfügung zu stellen, welches den Einsatz geringerer Lösungsmittelmengen erlaubt und im wesentlichen ohne Abfälle auskommt und zudem eine hohe Reinheit und höhere Ausbeuten erlaubt.

In diesem Zusammenhang soll ein möglichst großer Anteil der im Fermenter vorliegenden Nährstoffe, sowohl Carotinoide als auch weitere sich in den Mikroorganismen befindende, verwertet werden.

Somit ist es auch Aufgabe der vorliegenden Erfindung ein Verfahren zur Herstellung eines Carotinoid-haltigen Nahrungsmittels bereitzustellen, wobei das Nahrungsmittel selbst den Bedarf an Carotinoiden ohne Zusätze deckt. Insbesondere soll der Nährstoffgehalt der nach dem Verfahren erhältlichen Nahrungsmittel gegenüber den bisher erhältlichen Nahrungsmitteln zumindest gleichwertig sein. Ferner soll das Verfahren die effiziente Verwertung der produzierten Carotinoide ermöglichen.

Diese Aufgabe wird durch ein Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderten Organismen der Gattung *Blakeslea* gelöst, umfassend

- (i) Transformation mindestens einer der Zellen,
- (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder in mehreren genetischen Merkmalen alle gleichartig verändert sind und diese genetische Veränderung zur Ausprägung bringen, und
- (iii) Selektion und Vermehrung der gentechnisch veränderten Zelle oder Zellen,
- (iv) Kultivierung der gentechnisch veränderten Zellen,

- (v) Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe.

5 Mit der erfindungsgemäßen Methode ist es möglich, Blakeslea gezielt und stabil genetisch zu verändern, um so Mycel aus Zellen mit einheitlichen Kernen zu gewinnen, das Carotinoide oder deren Vorstufen, insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin produziert. Vorzugsweise handelt es sich um Zellen von
10 Pilzen der Art *Blakeslea trispora*. Die produzierten Carotinoiden oder deren Vorstufen sind dabei im wesentlichen frei von Verunreinigungen erhältlich und es können hohe Konzentrationen der Carotinoiden oder deren Vorstufen im Kulturmedium erzielt werden.

15 Unter Transformation wird die Übertragung einer genetischen Information in den Organismus, insbesondere Pilz verstanden. Darunter sollen alle dem Fachmann bekannten Möglichkeiten zur Einschleusung der Information, insbesondere DNA fallen, z. B. Beschuss mit DNA-beladenen Partikeln, Transformation mittels Protoplasten, Mikroinjektion von DNA, Elektroporation, Konjugation oder Transformation kompetenter Zellen, Chemikalien oder Agrobakterien vermittelte Transformation. Als genetische Information werden ein Genabschnitt, ein Gen oder mehrere Gene verstanden. Die genetische Information kann z. B. mit Hilfe eines Vectors oder als freie Nukleinsäure (z. B. DNA, RNA) und auf sonstige Weise in die Zellen
20 eingebracht und entweder durch Rekombination ins Wirtsgenom eingebaut oder in freier Form in der Zelle vorliegen. Besonders bevorzugt ist hierbei die homologe Rekombination.

Bevorzugte Transformationsmethode ist die *Agrobacterium tumefaciens*-
30 vermittelte Transformation. Hierzu wird zunächst die zu transferierende

Spender-DNA in einen Vektor eingefügt, der (i) flankierend zu der zu transferierenden DNA die T-DNA-Enden trägt, der (ii) einen Selektionsmarker enthält und der (iii) ggf. Promotoren und Terminatoren für die Genexpression der Spender-DNA aufweist. Dieser Vektor wird in einen Agrobacterium-tumefaciens-Stamm übertragen, der ein Ti-Plasmid mit den vir-Genen enthält. vir-Gene sind für den DNA-Transfer in Blakeslea verantwortlich. Mit diesem Zwei-Vektor-System wird die DNA von Agrobacterium in Blakeslea übertragen. Hierzu werden die Agrobakterien zunächst in Gegenwart von Acetosyringone inkubiert. Acetosyringone induziert die vir-Gene. Anschließend werden Sporen von Blakeslea trispora zusammen mit den induzierten Zellen von Agrobacterium tumefaciens auf Acetosyringone-haltigem Medium inkubiert und dann auf Medium übertragen, das eine Selektion der Transformanten, d.h. der gentechnisch veränderten Stämme von Blakeslea ermöglicht.

15

Der Begriff Vector wird in der vorliegenden Anmeldung als eine Bezeichnung für ein DNA-Molekül verwendet, das zum Einschleusen und ggf. zur Vermehrung von Fremd-DNA in eine Zelle dient (siehe auch "Vector" in Römpp Lexikon Chemie – CDRom Version 2.0, Stuttgart/New York: Georg Thieme Verlag 1999). In der vorliegenden Anmeldung sollen unter dem begriff "Vector" auch Plasmide, Cosmide usw. verstanden werden, die dem gleichen Zweck dienen.

20

Unter Expression wird in der vorliegenden Anmeldung die Übertragung einer genetischen Information ausgehend von DNA oder RNA in ein Genprodukt (hier vorzugsweise Enzyme zur Herstellung von Carotinoiden und insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin) verstanden und soll auch den Begriff der Überexpression beinhalten, womit eine verstärkte Expression gemeint ist, so dass ein bereits in der nicht transformierten Zelle (Wildtyp) hergestell-

25

30

tes Genprodukt verstärkt produziert wird oder einen großen Teil des gesamten Gehaltes der Zelle ausmacht.

5 Unter gentechnische Veränderung soll die Einschleusung genetischer Information in einen Empfängerorganismus, so dass diese stabil exprimiert und bei der Zellteilung weitergegeben wird, verstanden werden. In diesem Zusammenhang ist die Homokaryotisierung, die Herstellung von Zellen, die nur einheitliche Kerne enthalten, d. h. Kerne mit gleichem genetischem Informationsgehalt.

10

Diese Homokaryotisierung ist nur notwendig, wenn die durch Transformation eingeführte genetische Information rezessiv vorliegt, d. h. nicht zur Ausprägung gelangt. Führt die Transformation aber zu einem dominanten Vorliegen der genetischen Information, d. h. wird sie ausgeprägt, so ist
15 eine Homokaryotisierung nicht unbedingt nötig.

Vorzugsweise wird zur Homokaryotisierung eine Selektion der einkernigen Sporen durchgeführt. Von Natur aus ist ein geringer Anteil der Sporen von *Blakeslea trispora* einkernig, so dass sich diese ggf. nach spezifischer
20 Markierung z. B. Färbung der Zellkerne aussortieren lassen. Dies wird bevorzugterweise mittels FACS (Fluorescence Activated Cell Sorting) anhand der geringeren Fluoreszenz der einkernigen Zellen durchgeführt.

Alternativ kann zur Homokaryotisierung zunächst eine Kernreduktion
25 durchgeführt werden. Hierzu kann ein mutagenes Agens eingesetzt werden, wobei es sich insbesondere um N-Methyl-N'-nitro-nitrosoguanidin (MNNG) handelt. Auch die Verwendung von energiereichen Strahlen, wie UV- oder Röntgen-Strahlen zur Kernreduktion ist möglich. Anschließend kann zur Selektion auf das FACS Verfahren oder rezessive Selektions-
30 marker zurückgegriffen werden.

Unter Selektion wird die Auswahl von Zellen verstanden, deren Kerne dieselbe genetische Information beinhalten, d. h. Zellen die die gleichen Eigenschaften aufweisen, wie Resistenzen oder die Herstellung bzw. vermehrte Herstellung eines Produktes. In der Selektion werden neben der

5 FACS Methode bevorzugt 5-Carbon-5-deazariboflavin (DARF) und Hygromycin (hyg) oder 5'-Fluororotat (FOA) und Uracil eingesetzt.

Der in der Transformation (i) eingesetzte Vector kann derart gestaltet sein, dass die im Vector enthaltene genetische Information in das Genom mindestens einer Zelle integriert wird. Dabei kann genetische Information in

10 der Zelle ausgeschaltet werden. Dies kann direkt, d. h. durch eine Deletion erfolgen. Es ist aber auch möglich, daß der in der Transformation (i) eingesetzte Vector derart ausgestaltet ist, dass die im Vector enthaltene genetische Information in der Zelle exprimiert wird, d. h. genetische Information

15 eingefügt wird, die im korrespondierenden Wildtyp nicht vorhanden ist oder die durch die Transformation verstärkt bzw. überexprimiert wird und deren Produkt das Gen ausschaltet. Die eingeführte genetische Information kann aber auch indirekt eine genetische Information in der Zelle ausschalten, z. B. durch Produktion eines Inhibitors.

20

Der eingesetzte Vector enthält genetische Informationen oder Teile der genetischen Information zur Herstellung von Carotinoiden oder deren Vorstufen, insbesondere Carotinen oder Xanthophyllen oder deren Vorstufen. Der eingesetzte Vector enthält vorzugsweise genetische Informationen zur

25 Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, β -Carotin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, Lutein, Bixin oder Phytoen. Ganz besonders bevorzugt enthält der Vector Informationen zur Herstellung von Bixin, Phytoen, Canthaxanthin, Astaxanthin oder Zeaxanthin.

30

Der Vector kann beliebige genetische Informationen zur genetischen Veränderungen von Organismen der Gattung *Blakeslea* enthalten.

Unter „genetischer Information“ werden vorzugsweise Nukleinsäuren verstanden, deren Einbringung in den Organismus der Gattung *Blakeslea* zu einer genetischen Veränderung in Organismen der Gattung *Blakeslea*, also beispielsweise zu einer Verursachung, Erhöhung oder Reduzierung von Enzymaktivitäten im Vergleich zum Ausgangsorganismus führen.

Der Vector kann beispielsweise genetische Information zur Herstellung lipophiler Substanzen enthalten wie z.B. Carotinoide und deren Vorstufen, Phospholipide, Triacylglyceride, Steroide, Wachse, fettlösliche Vitamine, Provitamine und Cofaktoren oder genetische Information zur Herstellung hydrophiler Substanzen wie z.B. Eiweiße, Aminosäuren, Nukleotide und wasserlösliche Vitaminen, Provitamine und Cofaktoren.

Bevorzugterweise enthält der eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder Xanthophyllen oder deren Vorstufen.

Bevorzugterweise enthält der Vektor genetische Information, die eine Lokalisierung der Carotinoidbiosynthese-Enzyme in dem Zellkompartiment bewirkt, in dem die Carotinoidbiosynthese stattfindet.

Besonders bevorzugt sind genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3- und 3'-Hydroxyechinenon, Lycopin, Lutein, β -Carotin, Phytoen und/oder Phytofluen. Ganz besonders bevorzugt sind genetische Informationen zur Herstellung von Phytoen, Bixin, Lycopin, Zeaxanthin, Canthaxanthin und/oder Astaxanthin.

Entsprechend werden in einer bevorzugten Variante der Erfindung Organismen hergestellt und kultiviert, die über eine erhöhte Syntheserate für Zwischenprodukte der Carotinoidbiosynthese verfügen und folglich eine erhöhte Produktivität für Endprodukte der Carotinoidbiosynthese aufweisen. Zur Erhöhung der Syntheserate für Zwischenprodukte der Carotinoidbiosynthese werden insbesondere die Aktivitäten der Enzyme 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A-Reduktase (HMG-CoA-Reduktase), Isopentenylpyrophosphat-Isomerase und Geranylpyrophosphatsynthase gesteigert.

10

Entsprechend werden in einer besonders bevorzugten Variante der Erfindung Organismen hergestellt und kultiviert, die gegenüber dem Wildtyp eine erhöhte HMG-CoA-Reduktase-Aktivität aufweisen.

15 Unter HMG-CoA-Reduktase-Aktivität wird die Enzymaktivität einer HMG-CoA-Reduktase (3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A-Reduktase) verstanden.

Unter einer HMG-CoA-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A in Mevalonat umzuwandeln.

20

Dementsprechend wird unter HMG-CoA-Reduktase-Aktivität die in einer bestimmten Zeit durch das Protein HMG-CoA-Reduktase umgesetzte Menge 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A bzw. gebildete Menge Mevalonat verstanden.

25

Bei einer erhöhten HMG-CoA-Reduktase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein HMG-CoA-Reduktase die umgesetzte Menge 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A bzw. die gebildete Menge Mevalonat erhöht.

30

Vorzugsweise beträgt diese Erhöhung der HMG-CoA-Reduktase-Aktivität mindestens 5%, weiter bevorzugt mindestens 20%, weiter bevorzugt mindestens 50%, weiter bevorzugt mindestens 100%, besonders bevorzugt mindestens 300%, noch bevorzugter mindestens 500%, insbesondere
5 mindestens 600% der HMG-CoA-Reduktase-Aktivität des Wildtyps.

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der HMG-CoA-Reduktase-Aktivität gegenüber dem Wildtyp durch eine Erhöhung der Genexpression einer Nukleinsäure codierend eine HMG-CoA-Reduktase.
10

In einer besonders bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens erfolgt die Erhöhung der Genexpression einer Nukleinsäure codierend eine HMG-CoA-Reduktase indem man ein Nukleinsäurekonstrukt, enthaltend eine Nukleinsäure codierend eine HMG-CoA-Reduktase
15 in den Organismus einbringt, deren Expression in dem Organismus, verglichen mit dem Wildtyp, einer reduzierten Regulation unterliegt.

Unter einer reduzierten Regulation verglichen mit dem Wildtyp, wird eine im Vergleich zum vorstehend definierten Wildtyp verringerte, vorzugsweise
20 keine Regulation auf Expressions- oder Proteinebene verstanden.

Die reduzierte Regulation kann vorzugsweise durch einen im Nukleinsäurekonstrukt mit der kodierenden Sequenz funktionell verknüpften Promotor erreicht werden, der in dem Organismus verglichen mit dem Wildtyp-
25 Promoter einer reduzierten Regulation unterliegt.

Beispielsweise unterliegen die Promotoren ptef1 aus *Blakeslea trispora* und pgpdA aus *Aspergillus nidulans* nur einer reduzierten Regulation und sind daher insbesondere als Promotoren bevorzugt.

Diese Promotoren zeigen eine annähernd konstitutive Expression in *Blakeslea trispora*, so dass die transkriptionelle Regulation nicht mehr über die Intermediate der Carotinoidbiosynthese abläuft.

- 5 Die reduzierte Regulation kann in einer weiteren bevorzugten Ausführungsform dadurch erreicht werden, dass man als Nukleinsäure codierend eine HMG-CoA-Reduktase eine Nukleinsäure verwendet, deren Expression in dem Organismus, verglichen mit der Organismus eigenen, orthologen Nukleinsäure, einer reduzierten Regulation unterliegt.

10

Besonders bevorzugt ist die Verwendung einer Nukleinsäure, die nur den katalytischen Bereich der HMG-CoA-Reduktase kodiert (trunkierte (t-)HMG-CoA-Reduktase). Die für die Regulation verantwortliche Membran-Domäne fehlt. Die verwendete Nukleinsäure unterliegt somit einer reduzierten Regulation und führt zu einer Erhöhung der Genexpression der

15

HMG-CoA-Reduktase.

In einer besonders bevorzugten Ausführungsform bringt man Nukleinsäuren in *Blakeslea trispora* ein, welche die Sequenz SEQ ID. NO. 75 enthalten.

20

Weitere Beispiele für HMG-CoA-Reduktasen und damit auch für die auf den katalytischen Bereich reduzierten t-HMG-CoA-Reduktasen bzw. die kodierenden Gene lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Homologievergleiche der Sequenzen aus Datenbanken mit der SEQ ID. NO. 75 leicht finden.

25

Weitere Beispiele für HMG-CoA-Reduktasen und damit auch für die auf den katalytischen Bereich reduzierten t-HMG-CoA-Reduktasen bzw. die kodierenden Gene lassen sich weiterhin beispielsweise ausgehend von

30

der Sequenz SEQ ID. NO. 75 aus verschiedenen Organismen deren genomische Sequenz nicht bekannt ist, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

- 5 In einer besonders bevorzugten Ausführungsform wird die reduzierte Regulation dadurch erreicht, dass man als Nukleinsäure codierend eine HMG-CoA-Reduktase eine Nukleinsäure verwendet, deren Expression in dem Organismus, verglichen mit der Organismus eigenen, orthologen Nukleinsäure, einer reduzierten Regulation unterliegt und einen Promotor
10 verwendet, der in dem Organismus, verglichen mit dem Wildtyp-Promoter einer reduzierten Regulation unterliegt.

Entsprechend wird in einer bevorzugten Variante der Erfindung durch die Transformation die Genexpression der Phytoendesaturase ausgeschaltet,
15 so dass das von den Organismen produzierte Phytoen gewonnen werden kann. Der in der Transformation (i) eingesetzte Vector umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für ein Fragment des Gens der Phytoendesaturase, insbesondere carB aus *Blakeslea trispora* mit der SEQ ID NO: 69.

20

Entsprechend wird in einer bevorzugten Variante der Erfindung durch Transformation die Genexpression der Lycopincyclase ausgeschaltet, so dass das von den Organismen produzierte Lycopin gewonnen werden kann. Der in der Transformation eingesetzte Vektor umfasst daher in einer
25 Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für ein Fragment des Gens der Lycopincyclase, insbesondere carR aus *Blakeslea trispora*.

- 30 In einer bevorzugten Ausführungsform werden die Organismen der Gattung *Blakeslea* beispielsweise dadurch in die Lage versetzt Xanthophylle,

wie beispielsweise Canthaxanthin, Zeaxanthin oder Astaxanthin herzustellen, Bixin oder Phytoen, indem in den genetisch veränderten Organismen der Gattung *Blakeslea* im Vergleich zum Wildtyp eine Hydroxylase-Aktivität und/oder Ketolase-Aktivität verursacht wird.

5

Der in der Transformation (i) eingesetzte Vector enthält also in einer weiteren, bevorzugten Variante der Erfindung genetische Informationen, die nach Expression eine Ketolase- und/oder Hydroxylase-Aktivität entfalten, so dass die Organismen Zeaxanthin oder Astaxanthin produzieren.

10

Unter Ketolase-Aktivität wird die Enzymaktivität einer Ketolase verstanden.

Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -Ionon-Ring von Carotinoiden eine Keto-Gruppe einzuführen.

20

Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Canthaxanthin umzuwandeln.

Dementsprechend wird unter Ketolase-Aktivität die in einer bestimmten Zeit durch das Protein Ketolase umgesetzte Menge β -Carotin bzw. gebildete Menge Canthaxanthin verstanden.

Unter dem Begriff "Wildtyp" wird erfindungsgemäß der entsprechende nicht genetisch veränderte Ausgangsorganismus der Gattung *Blakeslea* verstanden.

30

Je nach Zusammenhang kann unter dem Begriff "Organismus" der Ausgangsorganismus (Wildtyp) der Gattung *Blakeslea* oder ein erfindungs-

gemäß, genetisch veränderter Organismus der Gattung *Blakeslea* oder beides verstanden werden.

Vorzugsweise wird unter "Wildtyp" für die Verursachung der Ketolase-Aktivität und für die Verursachung der Hydroxylase-Aktivität jeweils eine Referenz Organismus verstanden.

Dieser Referenzorganismus der Gattung *Blakeslea* ist *Blakeslea trispora* ATCC 14271 oder ATCC 14272, die sich lediglich im Paarungstyp unterscheiden.

Die Bestimmung der Ketolase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen der Gattung *Blakeslea* und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

Die Bestimmung der Ketolase-Aktivität in Organismen der Gattung *Blakeslea* erfolgt in Anlehnung an die Methode von Frazer et al., (J. Biol. Chem. 272(10): 6128-6135, 1997). Die Ketolase-Aktivität in Extrakten wird mit den Substraten beta-Carotin und Canthaxanthin in Gegenwart von Lipid (Sojalecithin) und Detergens (Natriumcholat) bestimmt. Substrat/Produkt-Verhältnisse aus den Ketolase-Assays werden mittels HPLC ermittelt.

Der erfindungsgemäße genetisch veränderte Organismus der Gattung *Blakeslea* weist in dieser, bevorzugten Ausführungsform im Vergleich zum genetisch nicht veränderten Wildtyp eine Ketolase-Aktivität auf und ist somit vorzugsweise in der Lage, transgen eine Ketolase zu exprimieren.

In einer weiter bevorzugten Ausführungsform erfolgt die Verursachung der Ketolase-Aktivität in den Organismen der Gattung *Blakeslea* durch Verursachung der Genexpression einer Nukleinsäure kodierend eine Ketolase.

In dieser bevorzugten Ausführungsform erfolgt die Verursachung der Genexpression einer Nukleinsäure kodierend eine Ketolase vorzugsweise durch Einbringen von Nukleinsäuren, die Ketolasen kodieren in den Ausgangsorganismus der Gattung *Blakesleaa*.

Dazu kann prinzipiell jedes Ketolase-Gen, also jede Nukleinsäure, die eine Ketolase codiert verwendet werden.

- 10 Alle in der Beschreibung erwähnten Nukleinsäuren können beispielsweise eine RNA-, DNA- oder cDNA-Sequenz sein.

Bei genomischen Ketolase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall, dass der Wirtsorganismus der Gattung *Blakesleaa* nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechenden Ketolase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu verwenden.

- 20 Beispiele für Nukleinsäuren, kodierend eine Ketolase und die entsprechenden Ketolasen, die im erfindungsgemäßen Verfahren verwendet werden können sind beispielsweise Sequenzen aus:

25 *Haematoccus pluvialis*, insbesondere aus *Haematoccus pluvialis* Flotow em. Wille (Accession NO: X86782; Nukleinsäure: SEQ ID NO: 11, Protein SEQ ID NO: 12),

Haematoccus pluvialis, NIES-144 (Accession NO: D45881; Nukleinsäure: SEQ ID NO: 13, Protein SEQ ID NO: 14),

Agrobacterium aurantiacum (Accession NO: D58420; Nukleinsäure: SEQ ID NO: 15, Protein SEQ ID NO: 16),

5 *Aliccaligenes spec.* (Accession NO: D58422; Nukleinsäure: SEQ ID NO: 17, Protein SEQ ID NO: 18),

Paracoccus marcusii (Accession NO: Y15112; Nukleinsäure: SEQ ID NO: 19, Protein SEQ ID NO: 20).

10 *Synechocystis sp.* Strain PC6803 (Accession NO: NP442491; Nukleinsäure: SEQ ID NO: 21, Protein SEQ ID NO: 22).

Bradyrhizobium sp. (Accession NO: AF218415; Nukleinsäure: SEQ ID NO: 23, Protein SEQ ID NO: 24).

15

Nostoc sp. Strain PCC7120 (Accession NO: AP003592, BAB74888; Nukleinsäure: SEQ ID NO: 25, Protein SEQ ID NO: 26),

Nostoc punctiforme ATTC 29133, Nukleinsäure: Acc.-No.

20 NZ_AABC01000195, Basenpaar 55,604 bis 55,392 (SEQ ID NO: 27); Protein: Acc.-No. ZP_00111258 (SEQ ID NO: 28) (als putatives Protein annotiert),

Nostoc punctiforme ATTC 29133, Nukleinsäure: Acc.-No.

25 NZ_AABC01000196, Basenpaar 140,571 bis 139,810 (SEQ ID NO: 29), Protein: (SEQ ID NO: 30) (nicht annotiert),

30 Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene, die im erfindungsgemäßen Verfahren verwendet werden können, lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Aminosäuresequenzen oder der

entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen und insbesondere mit den Sequenzen SEQ ID NO: 12, 26 und/oder 33 leicht auffinden.

- 5 Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 12, 26 und/oder 30 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich be-
- 10 kannter Weise leicht auffinden.

Die Hybridisierung kann unter moderaten (geringe Stringenz) oder vorzugsweise unter stringenten (hohe Stringenz) Bedingungen erfolgen.

- 15 Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 beschrieben.

20

Beispielhaft können die Bedingungen während des Waschschrilles ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer Stringenz (mit 2X SSC bei 50°C) und solchen mit hoher Stringenz (mit 0.2X SSC bei 50°C, bevorzugt bei 65°C) (20X SSC: 0,3 M Natriumcitrat, 3 M Natriumchlorid, pH 7.0).

25

Darüberhinaus kann die Temperatur während des Waschschrilles von moderaten Bedingungen bei Raumtemperatur, 22°C, bis zu stringenten Bedingungen bei 65°C angehoben werden.

30

Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS eingesetzt werden. In Gegenwart von 50% Formamid wird die Hybridisierung bevorzugt bei 42°C ausgeführt.

Einige beispielhafte Bedingungen für Hybridisierung und Waschschrift sind infolge gegeben:

10

- (1) Hybridisierungsbedingungen mit zum Beispiel
 - (i) 4X SSC bei 65°C, oder
 - (ii) 6X SSC bei 45°C, oder
 - (iii) 6X SSC bei 68°C, 100 mg/ml denaturierter Fischsperma-DNA, oder
 - 15 (iv) 6X SSC, 0.5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachs-sperma-DNA bei 68°C, oder
 - (v) 6XSSC, 0.5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachs-sperma-DNA, 50 % Formamid bei 42°C, oder
 - (vi) 50 % Formamid, 4X SSC bei 42°C, oder
 - 20 (vii) 50 % (vol/vol) Formamid, 0.1 % Rinderserumalbumin, 0.1 % Ficoll, 0.1 % Polyvinylpyrrolidon, 50 mM Natriumphosphatpuffer pH 6.5, 750 mM NaCl, 75 mM Natriumcitrat bei 42°C, oder
 - (viii) 2X oder 4X SSC bei 50°C (moderate Bedingungen), oder
 - (ix) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42°C (moderate Be-
25 dingungen).

- (2) Waschschriffe für jeweils 10 Minuten mit zum Beispiel
 - (i) 0.015 M NaCl/0.0015 M Natriumcitrat/0.1 % SDS bei 50°C, oder
 - (ii) 0.1X SSC bei 65°C, oder
 - 30 (iii) 0.1X SSC, 0.5 % SDS bei 68°C, oder
 - (iv) 0.1X SSC, 0.5 % SDS, 50 % Formamid bei 42°C, oder

- (v) 0.2X SSC, 0.1 % SDS bei 42°C, oder
- (vi) 2X SSC bei 65°C (moderate Bedingungen).

In einer bevorzugten Ausführungsform der erfindungsgemäßen genetisch
5 veränderten Organismen der Gattung *Blakeslea* bringt man Nukleinsäuren
ein, die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID
NO: 12 oder eine von dieser Sequenz durch Substitution, Insertion oder
Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von
mindestens 20 %, vorzugsweise mindestens 30%, 40%, 50%, 60%, be-
vorzugt mindestens 70%, 80%, besonders bevorzugt mindestens 90%,
10 insbesondere 91%, 92%; 93%, 94%, 95%, 96%, 97%, 98% oder 99% auf
Aminosäureebene mit der Sequenz SEQ ID NO: 12 und die enzymatische
Eigenschaft einer Ketolase aufweist.

15 Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die wie
vorstehend beschrieben durch Identitätsvergleich der Sequenzen aus an-
deren Organismen gefunden werden kann oder um eine künstliche Keto-
lase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 12 durch
künstliche Variation, beispielsweise durch Substitution, Insertion oder De-
20 letion von Aminosäuren abgewandelt wurde.

In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen
Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthal-
tend die Aminosäuresequenz SEQ ID NO: 26 oder eine von dieser Se-
25 quenz durch Substitution, Insertion oder Deletion von Aminosäuren abge-
leitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise
mindestens 30%, 40%, 50%, 60%, bevorzugt mindestens 70%, 80%, be-
sonders bevorzugt mindestens 90%, insbesondere 91%, 92%; 93%, 94%,
95%, 96%, 97%, 98% oder 99% auf Aminosäureebene mit der Sequenz
30 SEQ ID NO: 26 und die enzymatische Eigenschaft einer Ketolase auf-
weist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 26 durch
5 künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen
10 Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 30 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, 40 %, 50 %, bevorzugt mindestens 60 %, 70 %, bevor-
15 zugter mindestens 80 %, 85 % besonders bevorzugt mindestens 90 %, insbesondere 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % auf Aminosäureebene mit der Sequenz SEQ ID NO 30 und die enzymatische Eigenschaft einer Ketolase aufweist.

20 Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 30 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder De-
25 letion von Aminosäuren abgewandelt wurde.

Unter dem Begriff "Substitution" ist in der Beschreibung der Austausch einer oder mehrerer Aminosäuren durch eine oder mehrere Aminosäuren zu verstehen. Bevorzugt werden sog. konservative Austausche durchge-
30 führt, bei denen die ersetzte Aminosäure eine ähnliche Eigenschaft hat

wie die ursprüngliche Aminosäure, beispielsweise Austausch von Glu durch Asp, Gln durch Asn, Val durch Ile, Leu durch Ile, Ser durch Thr.

Deletion ist das Ersetzen einer Aminosäure durch eine direkte Bindung.

- 5 Bevorzugte Positionen für Deletionen sind die Termini des Polypeptides und die Verknüpfungen zwischen den einzelnen Proteindomänen.

Insertionen sind Einfügungen von Aminosäuren in die Polypeptidkette, wobei formal eine direkte Bindung durch ein oder mehrere Aminosäuren

- 10 ersetzt wird.

Unter Identität zwischen zwei Proteinen wird die Identität der Aminosäuren über die jeweils gesamte Proteinelänge verstanden, insbesondere die Identität die durch Vergleich mit Hilfe der Lasergene Software der Firma

- 15 DNASTAR, inc. Madison, Wisconsin (USA) unter Anwendung der Clustal Methode (Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl. Biosci. 1989 Apr;5(2):151-1) unter Einstellung folgender Parameter berechnet wird:

- 20 Multiple alignment parameter:

Gap penalty 10

Gap length penalty 10

Pairwise alignment parameter:

K-tuple 1

- 25 Gap penalty 3

Window 5

Diagonals saved 5

- 30 Unter einem Protein, das eine Identität von mindestens 20% auf Aminosäureebene mit der Sequenz SEQ ID NO: 12 oder 26 oder 30 aufweist, wird dementsprechend ein Protein verstanden, das bei einem Vergleich

seiner Sequenz mit der Sequenz SEQ ID NO: 12 oder 26 oder 30, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 20 %, bevorzugt 30%, 40%, 50%, besonders bevorzugt 60%, 70%, 80%, insbesondere 85%, 90, 95% aufweist.

5

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

Bevorzugt werden dafür solche Codons verwendet, die entsprechend der
10 Blakesleaspezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene von Organismen der Gattung Blakesleaa leicht ermitteln.

In einer besonders bevorzugten Ausführungsform bringt man eine Nu-
15 kleinsäure, enthaltend die Sequenz SEQ ID NO: 11 in die Organismus der Gattung ein.

In einer weiteren, besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 25 in die Orga-
20 nismus der Gattung ein.

In einer weiteren, besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 29 in die Orga-
nismus der Gattung ein.

25

Alle vorstehend erwähnten Ketolase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die
30 chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage,

Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Der in der Transformation (i) eingesetzte Vector umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für eine Ketolase, insbesondere der Ketolase *Nostoc punctiforme* aus mit der SEQ ID NO: 72.

Unter Hydroxylase-Aktivität die Enzymaktivität einer Hydroxylase verstanden.

Unter einer Hydroxylase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -Ionon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.

Insbesondere wird unter einer Hydroxylase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Zeaxanthin oder Cantaxanthin in Astaxanthin umzuwandeln.

Dementsprechend wird unter Hydroxyase-Aktivität die in einer bestimmten Zeit durch das Protein Hydroxylase umgesetzte Menge β -Carotin oder Cantaxanthin bzw. gebildete Menge Zeaxanthin oder Astaxanthin verstanden.

Bei einer erhöhten Hydroxylase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein

Hydroxylase die umgesetzte Menge β -Carotin oder Cantaxantin bzw. die gebildete Menge Zeaxanthin oder Astaxanthin erhöht.

Vorzugsweise beträgt diese Erhöhung der Hydroxylase-Aktivität mindestens 5 %, weiter bevorzugt mindestens 20 %, weiter bevorzugt mindestens 50 %, weiter bevorzugt mindestens 100 %, bevorzugter mindestens 300 %, noch bevorzugter mindestens 500 %, insbesondere mindestens 600 % der Hydroxylase-Aktivität des Wildtyps.

Die Bestimmung der Hydroxylase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

Die Aktivität der Hydroxylase wird nach Bouvier et al. (Biochim. Biophys. Acta 1391 (1998), 320-328) *in vitro* bestimmt. Es wird zu einer bestimmten Menge an Organismenextrakt Ferredoxin, Ferredoxin-NADP Oxidoreductase, Katalase, NADPH sowie beta-Carotin mit Mono- und Digalaktosylglyzeriden zugegeben.

Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, Keller, d'Harlingue und Camara (Xanthophyll biosynthesis: molecular and functional characterization of carotenoid hydroxylases from pepper fruits (Capsicum annuum L.; Biochim. Biophys. Acta 1391 (1998), 320-328):

Der in-vitro Assay wird in einem Volumen von 0.250 ml Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7.6), 0.025 mg Ferredoxin von Spinat, 0.5 Einheiten Ferredoxin-NADP⁺ Oxidoreduktase von Spinat, 0.25 mM NADPH, 0.010 mg beta-Carotin (in 0.1 mg Tween 80 emulgiert), 0.05 mM einer Mischung von Mono- und Digalaktosylglyzeriden (1:1), 1 Einheit Katalase, 200 Mono- und Digalaktosylglyzeriden, (1:1),

0.2 mg Rinderserumalbumin und Organismenextrakt in unterschiedlichem Volumen. Die Reaktionsmischung wird 2 Stunden bei 30°C inkubiert. Die Reaktionsprodukte werden mit organischem Lösungsmittel wie Aceton oder Chloroform/Methanol (2:1) extrahiert und mittels HPLC bestimmt.

5

Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, d'Harlingue und Camara (Molecular Analysis of carotenoid cyclase inhibition; Arch. Biochem. Biophys. 346(1) (1997) 53-64):

10

Der in-vitro Assay wird in einem Volumen von 250 µl Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7.6), unterschiedliche Mengen an Organismenextrakt, 20 nM Lycopin, 250 µg an chromoplastidärem Stromaprotein aus Paprika, 0.2 mM NADP⁺, 0.2 mM NADPH und 1 mM ATP. NADP/NADPH und ATP werden in 10 ml Ethanol mit 1 mg Tween 80 unmittelbar vor der Zugabe zum Inkubationsmedium gelöst. Nach einer Reaktionszeit von 60 Minuten bei 30°C wird die Reaktion durch Zugabe von Chloroform/Methanol (2:1) beendet. Die in Chloroform extrahierten Reaktionsprodukte werden mittels HPLC analysiert.

15

20

Ein alternativer Assay mit radioaktivem Substrat ist beschrieben in Fraser und Sandmann (Biochem. Biophys. Res. Comm. 185(1) (1992) 9-15).

25

Die Erhöhung der Hydroxylase-Aktivität kann durch verschiedene Wege erfolgen, beispielsweise durch Ausschalten von hemmenden Regulationsmechanismen auf Expressions- und Proteinebene oder durch Erhöhung der Genexpression von Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp.

30

Die Erhöhung der Genexpression der Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp kann ebenfalls durch verschiedene

Wege erfolgen, beispielsweise durch Induzierung des Hydroxylase-Gens durch Aktivatoren oder durch Einbringen von einer oder mehrerer Hydroxylase-Genkopien, also durch Einbringen mindestens einer Nukleinsäure kodierend eine Hydroxylase in den Organismus der Gattung *Blakesleaa*.

5

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der Genexpression einer Nukleinsäure kodierend eine Hydroxylase durch Einbringen von mindestens einer Nukleinsäure kodierend eine Hydroxylase in den Organismus der Gattung *Blakesleaa*.

10

Dazu kann prinzipiell jedes Hydroxylase-Gen, also jede Nukleinsäure, die eine Hydroxylase und jede Nukleinsäure, die eine β -Cyclase codiert, verwendet werden.

15 Bei genomischen Hydroxylase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall das der Wirtsorganismus nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechende Hydroxylase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu verwenden.

20

Ein Beispiel für ein Hydroxylase-Gen ist eine Nukleinsäure, kodierend eine Hydroxylase aus *Haematococcus pluvialis* mit der Accession No. AX038729 (WO 0061764; Nukleinsäure: SEQ ID NO: 31, Protein: SEQ ID NO: 32), aus *Erwinia uredovora* 20D3 (ATCC 19321, Accession No. D90087; Nukleinsäure: SEQ ID NO: 33, Protein: SEQ ID NO: 34) oder Hydroxylase aus *Thermus thermophilus* (DE 102 34 126.5) kodiert durch die Sequenz mit der SEQ ID NO 76.

25

sowie Hydroxylasen der folgenden Accession Nummern:
30 |emb|CAB55626.1, CAA70427.1, CAA70888.1, CAB55625.1, AF499108_1, AF315289_1, AF296158_1, AAC49443.1, NP_194300.1,

NP_200070.1, AAG10430.1, CAC06712.1, AAM88619.1, CAC95130.1,
AAL80006.1, AF162276_1, AAO53295.1, AAN85601.1, CRTZ_ERWHE,
CRTZ_PANAN, BAB79605.1, CRTZ_ALCSP, CRTZ_AGRAU,
CAB56060.1, ZP_00094836.1, AAC44852.1, BAC77670.1, NP_745389.1,
5 NP_344225.1, NP_849490.1, ZP_00087019.1, NP_503072.1,
NP_852012.1, NP_115929.1, ZP_00013255.1

In den erfindungsgemäßen bevorzugten transgenen Organismen der Gat-
tung *Blakeslea* liegt also in dieser bevorzugten Ausführungsform gegen-
10 über dem Wildtyp mindestens ein weiteres Hydroxylase-Gen vor.

In dieser bevorzugten Ausführungsform weist die genetisch veränderte
Organismus beispielsweise mindestens eine exogene Nukleinsäure, ko-
dierend eine Hydroxylase oder mindestens zwei endogene Nukleinsäuren,
15 kodierend eine Hydroxylase auf.

Bevorzugt verwendet man in vorstehend beschriebener bevorzugter Aus-
führungsform als Hydroxylase-Gene Nukleinsäuren, die Proteine kodieren,
enthaltend die Aminosäuresequenz SEQ ID NO: 32, 34 oder kodiert durch
20 die Sequenz mit der SEQ ID NO 76 oder eine von dieser Sequenz durch
Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Se-
quenz, die eine Identität von mindestens 30 %, vorzugsweise mindestens
50 %, bevorzugter mindestens 70%, noch bevorzugter mindestens 80 %,
am bevorzugtesten mindestens 90%, insbesondere 91%, 92%, 93%, 94%,
25 95%, 96%, 97%, 98%, 99% auf Aminosäureebene mit der Sequenz
SEQ. ID. NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO
76 und die die enzymatische Eigenschaft einer Hydroxylase aufweisen.

Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich
30 beispielsweise aus verschiedenen Organismen, deren genomische Se-
quenz bekannt ist, wie vorstehend beschrieben, durch Homologieverglei-

che der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit der SEQ ID. NO: 31, 33 oder 76 leicht auffinden.

- 5 Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich weiterhin beispielsweise ausgehend von der Sequenz SEQ ID NO: 31, 33 oder 76 aus verschiedenen Organismen deren genomische Sequenz nicht bekannt ist, wie vorstehend beschrieben, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

10

In einer weiter besonders bevorzugten Ausführungsform werden zur Erhöhung der Hydroxylase-Aktivität Nukleinsäuren in Organismen eingebracht, die Proteine kodieren, enthaltend die Aminosäuresequenz der Hydroxylase der Sequenz SEQ ID NO: 32, 34 oder kodiert durch die Sequenz mit
15 der SEQ ID NO 76.

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

- 20 Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Organismenspezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene der betreffenden Organismen leicht ermitteln.

- 25 In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ. ID. NO: 31, 33 oder 76 in den Organismus ein.

- Alle vorstehend erwähnten Hydroxylase-Gene sind weiterhin in an sich
30 bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender,

komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

10

Der in der Transformation (i) eingesetzte Vector umfasst daher in weiteren Ausführungsformen der Erfindung bevorzugterweise eine Sequenz codierend für eine Hydroxylase, insbesondere eine Hydroxylase aus *Haemotococcus pluvialis* mit der SEQ ID NO: 70 oder eine Hydroxylase aus *Erwinia uredova* mit der SEQ ID NO: 71. oder eine Hydroxylase aus *Thermus thermophilus* kodiert durch die Sequenz mit der SEQ ID NO 76.

15

Vorzugsweise wird durch die Transformation das Gen der Phytoendesaturase ausgeschaltet.

20

Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise ferner die Expression regelnde und unterstützende Bereiche, insbesondere Promotoren und Terminatoren.

25

Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise den *gpd* und/oder den *ptef1* Promotor und/oder den *trpC* Terminator. Diese haben sich zur Transformation der *Blakeslea* besonders bewährt. Auch der Einsatz von dem Fachmann geläufigen "inverted repeats" (IR, Römpp Lexikon der Biotechnologie 1992, Thieme Verlag Stuttgart, Seite 407 "Invers repetitive Sequenzen") zur Regelung der Expression bzw. Transkription liegt im Rahmen der Erfindung.

30

Vorteilhafterweise weist der im Vector eingesetzte gpd Promotor die Sequenz SEQ ID NO: 1 auf. Vorteilhafterweise weist der im Vector eingesetzte trpC Terminator die Sequenz SEQ ID NO: 2 auf.
5 Vorteilhafterweise weist der im Vector eingesetzte ptef1 Promotor die Sequenz SEQ ID NO: 35 auf.

Insbesondere werden dabei der gpd Promotor und der trpC Terminator aus *Aspergillus nidulans* und der ptef1 Promotor aus *Blakeslea trispora*
10 eingesetzt.

Insbesondere enthält der in der Transformation (i) eingesetzte Vector ein Resistenzgen. Bevorzugterweise handelt es sich um ein Hygromycin-Resistenzgen (hph), insbesondere eines aus *E. coli*. Dieses Resistenzgen
15 hat sich bei dem Nachweis der Transformation und Selektion der Zellen als besonders geeignet herausgestellt.

Als Promotor für hph wird also bevorzugt p-gpdA, der Promotor der Glycerinaldehyd-3-phosphatdehydrogenase aus *Aspergillus nidulans* genutzt.
20 Als Terminator für hph wird bevorzugt t-trpC, der Terminator des Gens trpC, codierend für Anthranilatsynthasekomponenten aus *Aspergillus nidulans* genutzt.

Als Vektoren haben sich Abkömmlinge des pBinAHyg Vectors als besonders geeignet herausgestellt. Der zur Transformation eingesetzte Vector umfasst also bevorzugterweise die SEQ ID NO: 3.

Hinzu kommen je nach gewünschtem Carotinoid oder dessen Vorstufe eine Sequenz codierend für eine Hydroxylase, Ketolase, Phytoendesaturase usw. wie diese zuvor beschrieben wurden. Die Vektoren umfassen
30 also in einer Ausführungsform der Erfindung die Sequenz SEQ ID NO: 69 codierend für die Phytoendesaturase. Die Vektoren umfassen ferner in

einer weiteren Ausführungsform der Erfindung die Sequenz SEQ ID NO: 72 codierend für eine Ketolase. Die Vektoren umfassen weiter in einer weiteren Ausführungsform der Erfindung die Sequenz SEQ ID NO: 70 oder 71 oder 76 codierend für eine Hydroxylase. Entsprechende Kombinationen
5 der zuvorgenannten Sequenzen liegen ebenso im Rahmen der Erfindung. So umfasst der Vector in einer Ausführungsform sowohl eine Sequenz SEQ ID NO: 72 codierend für eine Ketolase als auch die Sequenz SEQ ID NO: 70 oder 71 oder 76 codierend für eine Hydroxylase und ermöglicht so die Herstellung von Astaxanthin.

10

Insbesondere sind Vektoren ausgewählt aus der Gruppe bestehend aus den SEQ ID NO: 37 bis 51 und 62 im Rahmen der Erfindung einsetzbar.

15

Die genetisch veränderten Organismen können zur Produktion von Carotinoiden, Xanthophyllen oder deren Vorstufen, insbesondere Bixin, Phytoen, Astaxanthin, Zeaxanthin und Canthaxanthin verwendet werden. Auch können neue, im Wildtyp natürlicherweise nicht vorkommende Carotinoide durch Einbringung der entsprechenden genetischen Information von den gezielt genetisch veränderten Zellen bzw. dem durch sie gebildeten Mycel erzeugt und anschließend isoliert werden.

20

Die gentechnisch veränderten Zellen werden nach der Selektion kultiviert, so daß Carotinoiden oder deren Vorstufen bereitgestellt werden können.

25

Bevorzugterweise ist die Gewinnung von Carotinoiden oder deren Vorstufen mit den gezielt genetisch veränderten Zellen bzw. das durch sie gebildete Mycel möglich.

30

Die Kultivierung der Organismen unterliegt keinen Besonderheiten. Vorteilhafterweise werden, insbesondere bei der Verwendung von *Blakeslea*

trisporea, entgegengesetzte Paarungstypen gemeinsam kultiviert, da dies zu besserem Wachstum und Produktion führt.

5 Wird die gentechnische Veränderung nur in Zellen eines der vorkommen-
den Paarungstypen (bei *Blakeslea trispora* (+) oder (-)) durchgeführt, so
wird zur Kultivierung der entsprechend andere, nicht veränderte Paarungs-
typ zugesetzt, da so eine gute Produktion der Carotinoide oder deren Vor-
stufen aufgrund der von dem zweiten, nicht veränderten Paarungstyp ab-
gegebenen Substanzen (z. B. Trisporsäuren) zu erreichen ist. Vorteilhaft-
10 erweise wird jedoch die gentechnische Veränderung in Zellen beider Paa-
rungstypen vorgenommen und diese zusammen kultiviert. Hierdurch wird
ein besonders gutes Wachstum und eine optimale Produktion der Caroti-
noiden oder deren Vorstufen erreicht. Auch eine (künstliche) Zugabe der
Trisporsäuren ist möglich und sinnvoll.

15

Trisporsäuren sind Sexualhormone in Mucorales Pilzen, wie *Blakeslea*,
welche die Bildung von Zygothoren und die Produktion von β -Carotin sti-
mulieren (van den Ende 1968, J. Bacteriol. 96:1298 - 1303, Austin et al.
1969, Nature 223:1178 - 1179, Reschke Tetrahedron Lett. 29:3435 -
20 3439, van den Ende 1970, J. Bacteriol. 101:423 - 428).

Es können alle dem Fachmann geläufigen Medien eingesetzt werden, so
weit sich diese zur Kultivierung der eingesetzten Organismen und deren
Carotinoid Produktion eignen. Insbesondere müssen bei Einsatz der
25 GVO keine Carotinoidbiosynthese Inhibitoren eingesetzt werden. Die ein-
gesetzten Medien beinhalten vorzugsweise Zusätze, wie eine oder mehre-
re Kohlenstoffquellen, eine oder mehrere Stickstoffquellen, Mineralsalze
und Thiamine. Bevorzugterweise werden Zusätze eingesetzt, wie sie aus
der WO 03/038064 A2, Seite 4, Zeile 30 bis Seite 5 Zeile 7 hervorgehen.
30 Besonders bevorzugt wird als Kohlenstoffquelle Glukose und als Stick-

stoffquelle Asparagin, pflanzliche oder tierische Extrakte, wie Baumwollsaatöl, Sojaöl, Baumwollsamemehl oder Hefe-Extrakt zugesetzt.

Die Kultivierung kann entweder unter aeroben oder anaeroben Bedingungen durchgeführt werden. Auch eine gemischte, zunächst aerobe und anschließend anaerobe Kultivierung, wie sie aus der DE 101 30 323 bekannt ist, ist möglich. Temperatur und Luftfeuchtigkeit werden dabei jeweils zum optimalen Wachstum eingestellt. Bevorzugterweise liegt die Temperatur bei der Kultivierung zwischen ca. 20 und ca. 34 °C, insbesondere zwischen ca. 26°C und ca. 28°C. Die Kultivierung kann ferner kontinuierlich, batch- oder satzweise erfolgen.

Die Kultivierung erfolgt vorzugsweise bis zu einem Feststoffgehalt zwischen etwa 1 und etwa 20 %, bevorzugt 3 und 15 % und besonders bevorzugt 4 und 11 %. Insbesondere ist wichtig dass die Kulturbrühe pumpbar bleibt, so dass sie in den nachfolgenden Verfahrensschritten ver- und bearbeitbar bleibt. Ist der Feststoffgehalt zu klein, so muss ein großer Aufwand bei der Aufkonzentrierung oder Trocknung betrieben werden.

Die Kultivierung bzw. Fermentation kann in den üblichen Apparaturen durchgeführt werden. Hierzu kommen alle für die jeweils eingesetzten Mikroorganismen und deren Produkte geeigneten Apparaturen in Betracht. Insbesondere solche, wie sie aus dem Römpp Lexikon Biotechnologie (1992 Georg Thieme Verlag, Stuttgart) unter dem Stichwort "Bioreaktor" auf Seiten 123 - 126 angegeben sind. Besonders bevorzugt ist der Einsatz von Rührkesselreaktoren mit versch. Einbauten, Blasensäulen verschiedener Bauarten, etc.

Die nach dem erfindungsgemäßen Verfahren bereitgestellten Carotinoiden oder deren Vorstufen, insbesondere Bixin, Phytoen oder Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin eignen sich besonders

zur Herstellung von Zusätzen für Futter-, Nahrungs- und Nahrungsergänzungsmittel, kosmetischen, pharmazeutischen oder dermatologischen Zubereitungen.

- 5 Die Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe aus der Kultur der gentechnisch veränderten Mikroorganismen erfolgt nach zwei Varianten a) oder b), bevorzugt ist auch eine Kombination aus a) und b);

10

a:

I) Abtrennung der Biomasse,

IA) ggf. Waschen der Biomasse mit einem Carotinoide nicht lösenden Lösungsmittel, insbesondere Wasser,

15

IB) Sterilisation und Zellaufschluß der Biomasse,

IC) ggf. Trocknung und/oder homogene Verteilung und

II) partielle Extraktion der Carotinoide aus der aufgeschlossenen Biomasse mittels eines Carotinoide lösenden Lösungsmittels und Trennung des Lösungsmittels von der Biomasse,

20

IIA)

1) Entfernung von Lösemittelresten aus der Carotinoidhaltigen Biomasse,

2) ggf. homogene Suspension der Biomasse mit einem Biomasse-Feststoffgehalt > 2 % und < 50 %, und

25

3) Trocknung der Biomasse bzw. Suspension zur Herstellung des Nahrungsmittels,

IIB)

1) Kristallisation der Carotinoide aus dem verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration;

30

oder b):

- 5 i) Homogene Suspendierung der Feststoffe der Kulturbrühe
 und
- IIA) bei einem Feststoffgehalt der Kulturbrühe von > 2 %
 1) ggf. Konzentration der Kulturbrühe auf einen Fest-
 stoffgehalt < 50 % und
10 2) Trocknung der Kulturbrühe zur Herstellung des Nah-
 rungsmittels
- oder
- IIB) bei einem Feststoffgehalt von < 2 % der Kulturbrühe,
15 1) Konzentration der Kulturbrühe auf einen Feststoffge-
 halt > 2 % und < 50 % und
 2) Trocknung der Suspension zur Herstellung des Nah-
 rungsmittels,
- oder
- 20 IIC) unabhängig vom Feststoffgehalt der Kulturbrühe,
 1) Abtrennung der Biomasse,
 2) ggf. Waschen der Biomasse mit Carotinoide nicht lö-
 sende Lösungsmitteln, insbesondere Wasser,
25 3) Sterilisation und Zellaufschluß,
 4) ggf. Trocknung und homogene Verteilung,
 5) partielle Extraktion der Carotinoide aus der Biomasse
 mittels eines Carotinoide lösendes Lösungsmittels,
 5a) Abtrennung der Carotinoid-haltigen Biomasse vom
30 Carotinoid-haltigen Lösungsmittel,

- 5b) Entfernung von Lösemittelresten aus der Biomasse und
- 5c) Trocknung der Biomasse zur Herstellung des Nahrungsmittels,
- 5 6) Kristallisation der Carotinoide aus dem in 5a) verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration.

10 Die erfindungsgemäße Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe aus der Kultur der gentechnisch veränderten Mikroorganismen erfolgt nach zwei Varianten a) oder b) ermöglicht die gleichzeitige Herstellung von zwei Produkten.

15

Durch die erfindungsgemäße Kombination der Herstellung von zwei Produkten, insbesondere bei der Bereitstellung gemäß Variante a), nämlich dem mindestens einem Carotinoid und dem Carotinoid-haltigem Nahrungsmittel ist keine vollständige Extraktion der Carotinoide aus der Bio-

20 masse nötig, so dass der Aufwand bei der Extraktion geringer ausfällt. Das Carotinoid muss trotz vollständiger Verwertung nur partiell extrahiert werden, ohne dass es zu Produktverlusten kommt. Dies bedingt geringere Lösungsmittelmengen und damit einhergehend einen geringeren Aufwand bei den Maßnahmen zu deren Wiederverwendung. Zudem werden Abfälle

25 weitestgehend vermieden, da die Biomasse nicht als Abfall anfällt, sondern zum hochwertigen Nahrungsmittel weiterverarbeitet wird. Somit ergeben sich geringere Kosten für die Verfahren durch Ausnutzen von Synergien.

30 Nach dem erfindungsgemäßen Verfahren mit Bereitstellung gemäß Variante b) erhältliche Nahrungsmittel enthalten also bereits nach der Herstel-

lung große Mengen an Carotinoiden, die nicht zugesetzt werden müssen. Dadurch, dass das Nahrungsmittel neben dem mindestens einen Carotinoid auch *Blakeslea trispora* enthält ist sein Nährstoffgehalt zudem gesteigert. Insbesondere ist der Nährstoffgehalt nach den bevorzugten Alternativen IIA und IIB stark gesteigert, da es neben dem mindestens einen Carotinoid und *Blakeslea trispora* zusätzlich alle Medienbestandteile der Fermentation enthält. Ferner benötigt das Verfahren keine zusätzlichen, aufwendigen Aufarbeitungs- und Herstellungsschritte, sondern die homogenisierte und ggf. entwässerte *Blakeslea trispora*-haltige Kulturbrühe kann direkt ohne Umwege zur Herstellung des Nahrungsmittels getrocknet werden. Es fallen demnach praktisch keine Abfälle an, abgesehen vom wässrigem Medium bei der Alternative IIB, welches jedoch unproblematisch in einer Kläranlage gereinigt werden kann. Zusätzlich wird in allen drei Alternativen die gesamte Produktionsmenge Carotinoide ohne oder mit nur marginalen Verlusten verwertet, da gemäß IIA und IIB keine verlustreichen Trenn- bzw. Aufarbeitungsschritte vorgenommen werden müssen. Bei der Alternative IIIC wird ebenfalls die gesamte Produktionsmenge Carotinoide ohne oder mit nur marginalen Verlusten verwertet, da ein Teil in der Biomasse zum Nahrungsmittel verarbeitet wird und der andere Teil zur Gewinnung reiner Carotinoide extrahiert wird. Durch die erfindungsgemäße Kombination der Herstellung von zwei Produkten gemäß IIC, nämlich dem Carotinoid-haltigem Nahrungsmittel und den Carotinoiden an sich, ist vorteilhaft, dass wiederum im wesentlichen keine Abfälle entstehen und eine vollständige Extraktion der Carotinoide aus der Biomasse unnötig ist, so dass sonst bei der Extraktion anfallende Aufwand geringer ausfällt. Das oder die wertvollen Carotinoide müssen trotz vollständiger Verwertung nur partiell extrahiert werden, ohne dass es zu Produktverlusten kommt. Dies bedingt geringere Lösungsmittelmengen und damit einhergehend einen geringeren Aufwand bei den Maßnahmen zu deren Wiederverwendung. Zudem werden Abfälle weitestgehend vermieden, da die Biomasse nicht als Abfall anfällt, sondern zum hochwertigen Nahrungsmittel verarbeitet

wird. Somit ergeben sich geringere Kosten für die Verfahren durch Ausnutzen von Synergien.

Unter "hochrein" soll in der vorliegenden Anmeldung eine Reinheit des
5 mindestens einen Carotinoids von mindestens 95%, bevorzugt > 95%,
vorzugsweise > 96%, besonders bevorzugt > 97%, ganz besonders be-
vorzugt > 98%, höchst bevorzugt > 99% verstanden werden.

Als nach dem erfindungsgemäßen Verfahren herstellbare Carotinoide
10 kommen alle natürlichen und künstlichen Carotine und Xanthophylle in
Betracht. Insbesondere ist das mindestens eine Carotinoid aus der Grup-
pe bestehend aus Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin,
Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon,
3'-Hydroxyechinenon, Lycopin, β -Carotin, Lutein, Phytofluen, Bixin und
15 Phytoen ausgewählt. Bevorzugterweise handelt es sich um Astaxanthin
oder Zeaxanthin. Die Carotinoide können nach dem erfindungsgemäßen
Verfahren einzeln oder als Gemische von zwei oder mehrerer der zuvor
genannten Carotinoide erhalten werden. Insbesondere bei Einsatz der
weiter unten angegebenen gentechnisch veränderten Organismen (GVO)
20 kann das oder können die Carotinoide gezielt hergestellt werden.

Als Nahrungsmittel werden Zusammensetzungen angesehen, die der Er-
nährung dienen. Darunter fallen auch Zusammensetzungen für die Ergän-
zung der Ernährung. Insbesondere werden als Nahrungsmittel Tierfutter-
25 mittel und Tierfutterergänzungsmittel angesehen.

Nach der Kultivierung kann, gemäß Variante a) der Bereitstellung, die
Biomasse von der Kulturbrühe abgetrennt. Hierzu können alle dem Fach-
mann geläufigen und üblicherweise einsetzbaren Methoden zur
30 fest/flüssig-Trennung eingesetzt werden. Hierunter fallen insbesondere die
mechanischen Verfahren, wie Filtration und Zentrifugation, die auf der

Ausnutzung von Schwerkraft, Zentrifugalkraft, Druck oder Vakuum beruhen. Zu den einsetzbaren Verfahren und Apparaten gehören daneben u. a. Querstromfiltration bzw. Membrantechniken wie Osmose, umgekehrte Osmose, Mikrofiltration, Ultrafiltration, Nanofiltration, Kuchenfiltrationsverfahren (z.B. mittels Pressfilterautomaten, (Membran-, Rahmen- oder Kammer-)Filterpressen, (Rühr-)drucknutschen, Saugnutschen, (Vakuum-)bandfiltern, (Vakuum-)Trommelfiltern, Drehfiltern, Kerzenfiltern), Zentrifugationsverfahren mittels kontinuierlich oder diskontinuierlich betriebener Zentrifugen oder Filterzentrifugen (z.B. StülpfILTERZentrifugen, Schälzentrifugen, Schubzentrifugen, Siebschneckenzentrifugen, Gleitzentrifugen, Separatoren oder Dekantern), Verfahren unter Ausnutzung der Schwerkraft wie Flotation, Sedimentation, Sink-Schwimm-Aufbereitung und Klären. Bevorzugterweise erfolgt die Abtrennung der Biomasse von der Kulturbrühe durch Zentrifugation mittels eines Dekantern oder durch Filtration, mittels einer Membranfiltrationseinheit durchgeführt.

In dem zweiten Schritt der Bereitstellung nach Variante b) wird eine homogen verteilte Suspension der Feststoffe in der Kulturbrühe erzeugt. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden verwendet werden. Insbesondere kommen dazu Dispergiergeräte, wie ein Ultra-Turrax® (im Labormaßstab) zum Einsatz. Ein Zellaufschluss ist nicht notwendig, kann aber vorgenommen werden.

Falls notwendig, kann die Kulturbrühe entwässert werden, um einen geeigneten Feststoffgehalt zwischen $> 2 \%$ und $< 50 \%$ zu erreichen. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden zur fest/flüssig-Trennung eingesetzt werden. Hierunter fallen insbesondere die mechanischen Verfahren, wie Filtration und Zentrifugation, die auf der Ausnutzung von Schwerkraft, Zentrifugalkraft, Druck oder Vakuum beruhen. Zu den einsetzbaren Verfahren und Apparaten gehören daneben u. a. Querstromfiltration bzw. Membrantechniken wie Osmose,

umgekehrte Osmose, Mikrofiltration, Ultrafiltration, Nanofiltration, Kuchenfiltrationsverfahren (z.B. mittels Pressfilterautomaten, (Membran-, Rahmen- oder Kammer-)Filterpressen, (Rühr-)drucknutschen, Saugnutschen, (Vakuum-)bandfiltern, (Vakuum-)Trommelfiltern, Drehfiltern, Kerzenfiltern),

5 Zentrifugationsverfahren mittels kontinuierlich oder diskontinuierlich betriebener Zentrifugen oder Filterzentrifugen (z.B. Stülpfilterzentrifugen, Schälzentrifugen, Schubzentrifugen, Siebschneckenzentrifugen, Gleitzentrifugen, Separatoren oder Dekantern), Verfahren unter Ausnutzung der Schwerkraft wie Flotation, Sedimentation, Sink-Schwimm-Aufbereitung

10 und Klären. Bevorzugterweise erfolgt die Abtrennung der Biomasse von der Kulturbrühe durch Zentrifugation mittels eines Dekantern oder wird durch Filtration, mittels einer Membranfiltrationseinheit durchgeführt. Anschließend wird die Kulturbrühe getrocknet. Hierzu können wiederum alle dem Fachmann bekannten Verfahren und Apparate eingesetzt werden.

15 Insbesondere eignen sich Apparate zur thermischen Trocknung wie Konvektions-, Kontakt- und Strahlungstrocknung, z.B. Horden-, Kammer-, Kanal-, Flachbahn-, Teller-, Drehtrommel-, Rieselschacht-, Siebband-, Strom-, , Wirbelschicht-, Fließbett-, Schaufel-, Kugelbett-, , Heizteller-, Dünnschicht-, Walzen-, Band-, Siebtrommel-, Schnecken-, Taumel-, Kontakt-Scheiben-, Infrarot-, Mikrowellen- und Gefriertrockner, Sprühtrockner

20 oder Sprühtrockner mit integrierter Wirbelschicht, die durch Dampf, Öl, Gas oder elektrischen Strom ggf. beheizt und ggf. unter Vakuum betrieben werden. Die Betriebsweise kann dabei je nach Apparat kontinuierlich oder diskontinuierlich sein. Daneben oder damit in Kombination können die

25 oben bereits angegeben mechanischen Verfahren zur Fest/flüssig-Trennung verwendet werden.

Eine Granulierung durch Extrusion wie dies aus der WO 97/36996 A2 hervorgeht ist jedoch nicht notwendig. Durch die Trocknung wird das Nahrungsmittel haltbar und lagerfähig.

Insbesondere wird die Kulturbrühe sprühgetrocknet. Bevorzugt wird zur Trocknung die Sprühtrocknung eingesetzt, wie sie aus der DE 101 04 494 A1, DE-A-12 11 911 oder EP 0 410 236 A1 bekannt sind. Ergänzend wird auf vgl. Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Sprühtrocknung" und Römpp Lexikon Biotechnologie, Georg Thieme Verlag, 1992, "Zerstäubungstrocknung" verwiesen. Die Sprühtrocknung bietet den Vorteil der kurzen Verweilzeit des Produkts in der heißen Zone des Trockners, so dass eine besonders schonende Trocknung erzielt wird.

10

Bei der Sprühtrocknung werden Eingangstemperaturen von ca. 115°C – 180°C, bevorzugt 120°C – 130°C, und Ausgangstemperaturen von ca. 50°C – 80°C, bevorzugt 55°C – 70°C gewählt. Als Trocknungsgas wird vorzugsweise Stickstoff eingesetzt.

15

Gegebenenfalls können zur Erzielung einer besseren Rieselfähigkeit Rieselfähigungsmittel, wie Kieselsäuren etc. zugesetzt werden. Der Einsatz von inerten Trägermaterialien, d.h. niedermolekularen anorganischen Trägern wie NaCl, CaCO₃, Na₂SO₄ oder MgSO₄, organischen Trägern wie Glucose, Fructose, Saccharose, Dextrine oder Stärkeprodukten (Roggen-, Gersten-, Hafermehl, Weizengrießkleie) ist denkbar.

20

Das getrocknete Produkt weist bevorzugt eine Restfeuchte von weniger als 10 %, bevorzugt weniger als 5 % bezogen auf die Trockenmasse auf. Sein Carotinoidgehalt liegt zwischen 0,05 und 20 %, insbesondere 1 und 10 % bezogen auf die Trockenmasse.

25

Das so hergestellte Nahrungsmittel kann entweder direkt verwendet werden oder mittels weiterer Zusätze aufbereitet werden, so wie dies ebenfalls aus der DE 101 04 494 A1 bekannt ist.

30

Gemäß der Alternative IIC wird nach der Kultivierung und vor der Trocknung der Biomasse, die Biomasse von der Kulturbrühe zunächst abgetrennt. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden zur fest/flüssig-Trennung eingesetzt werden, wie
5 sie bereits oben bei der Entwässerung genannt wurden. Bevorzugterweise erfolgt die Abtrennung der Biomasse von der Kulturbrühe durch Zentrifugation mittels eines Dekanters oder durch Membranfiltration durchgeführt.

Anschließend erfolgt optional das Waschen der Biomasse mit einem Carotinoide nicht lösenden Lösungsmittel, insbesondere Wasser, wodurch insbesondere Wasser lösliche Komponenten entfernt werden. Dieser Schritt kann gegebenenfalls unter Verwendung weiterer Carotinoide nicht lösenden Lösungsmittel (z. B. Alkohole) ergänzt werden, was aber im Rahmen der Erfindung nicht notwendig ist und zur Vermeidung von Abfällen nicht
15 bevorzugt ist.

Anschließend erfolgen die Sterilisation und der sich anschließende oder gleichzeitige Zellaufschluß der Zellen in der Biomasse. Durch die Sterilisation werden die Mikroorganismen abgetötet und gegebenenfalls vorhandene Enzymaktivität beendet. Dies ist zur Verhinderung des Abbaus der
20 Biomasse bzw. der darin enthaltenen Stoffe, insbesondere der Carotinoide und für die Haltbarkeit von Bedeutung.

Die Sterilisation kann mit einem üblichen, dem Fachmann geläufigen Verfahren durchgeführt werden. Hierzu gehören die Sterilisation mittels Dampf, insbesondere bei Temperaturen größer 120 °C unter Druck (≥ 1 bar) und Zeitauern von \geq ca. 20 min. sowie die Behandlung mit energiereichen Strahlen, wie UV-, Mikrowellen, Gamma- oder Beta-Strahlen. Bevorzugterweise erfolgt die Sterilisation im Rahmen des erfindungsgemäßen Verfahrens mittels Dampf oder Mikrowellenstrahlung.
30

Durch den nachfolgenden oder gleichzeitigen Zellaufschluß, werden die innerhalb der Zellen vorliegenden Carotinoide freigesetzt. Der Zellaufschluß kann ebenfalls mit allen dem Fachmann bekannten üblichen Verfahren erfolgen. Hierzu gehören mechanische und nicht mechanische Methoden. Zu den mechanischen Methoden zählen Trockenmahlen, Naßmahlen, Rühren, Homogenisieren (z.B. im Hochdruckhomogenisator) und die Verwendung von Ultraschall oder Mikrowellen. Als nicht mechanische Methoden kommen physikalische, chemische und biochemische Methoden in Betracht. Hierzu gehören Kurzzeiterhitzen, Kurzzeitgefrieren, Osmotischer Schock, Trocknung, Behandlung mit Säuren oder Laugen sowie ein enzymatischer Aufschluss. Günstiger Weise wird zum Zellaufschluss jedoch das zur Sterilisation verwendete Verfahren eingesetzt. Bevorzugterweise wird also ebenfalls der Zellaufschluss mittels Dampf oder Mikrowellen Strahlung durchgeführt.

15

Die Sterilisation und/oder der Zellaufschluss können kontinuierlich oder diskontinuierlich durchgeführt werden.

Die Sterilisation und/oder der Zellaufschluss können im zur Kultivierung eingesetzten Bioreaktor oder in anderen Apparaturen, wie Autoklaven usw. durchgeführt werden. Bei kontinuierlicher Durchführung kann das aus der WO 01/83437 A1 bekannte Mikrowellen verwendende Verfahren und entsprechende Apparaturen eingesetzt werden.

25 Vor der Extraktion wird die Biomasse gegebenenfalls getrocknet und/oder homogenisiert. Hierzu können wiederum alle dem Fachmann bekannten üblichen Verfahren und Geräte eingesetzt werden. Insbesondere eignen sich Apparate zur thermischen Trocknung wie Konvektions-, Kontakt- und Strahlungstrocknung, z.B. Horden-, Kammer-, Kanal-, Flachbahn-, Teller-,
30 Drehtrommel-, Rieselschacht-, Siebband-, Strom-, , Wirbelschicht-, Fließbett-, Schaufel-, Kugelbett-, , Heizteller-, Dünnschicht-, Walzen-, Band-,

Siebtrommel-, Schnecken-, Taumel-, Kontakt-Scheiben-, Infrarot-, Mikrowellen- und Gefriertrockner, Sprühtrockner oder Sprühtrockner mit integrierter Wirbelschicht, die durch Dampf, Öl, Gas oder elektrischen Strom ggf. beheizt und ggf. unter Vakuum betrieben werden. Die Betriebsweise
5 kann dabei je nach Apparat kontinuierlich oder diskontinuierlich sein. Daneben oder damit in Kombination können die oben bereits angegeben mechanischen Verfahren zur Fest/flüssig-Trennung verwendet werden.

Eine Granulierung durch Extrusion wie dies aus der WO 97/36996 A2 hervorgeht ist jedoch nicht notwendig.
10

Anschließend erfolgt die partielle Extraktion der Carotinoide aus der aufgeschlossenen Biomasse mittels eines Carotinoide lösenden Lösungsmittels und Trennung des Lösungsmittels von der Biomasse. Sowohl in dem
15 Lösungsmittel als auch in der Biomasse sind nun Carotinoide enthalten, wobei sich in dem Lösungsmittel bevorzugterweise der Großteil der Carotinoide befindet.

Aus dem Lösungsmittel werden anschließend die hochreinen Carotinoide
20 isoliert, wohingegen die Biomasse zu einem hochwertigen, Carotinoidhaltigen Nahrungsmittel weiterverarbeitet wird, welches durch den vorhergehenden Zellaufschluss auch eine gute Bioverfügbarkeit der Carotinoide aufweist.

25 Unter partieller Extraktion soll demnach die bewusst unvollständige Extraktion der Carotinoiden aus der Biomasse verstanden werden (vgl. oben). Bevorzugterweise wird im Rahmen der Erfindung durch die Extraktion also weniger als 100 % der in der Biomasse enthaltenen Gesamtmenge der Carotinoide aus dieser extrahiert.

30 Dies ist von großem Vorteil, da der Aufwand zur Extraktion mit der abnehmenden Menge Carotinoid in der Biomasse überproportional zunimmt.

Zur Extraktion werden Lösungsmittel eingesetzt, die Carotinoide lösen, wie z. B. Hexan, Ethylacetat, Dichlormethan oder überkritisches Kohlendioxid. Bevorzugterweise wird erfindungsgemäß als Lösungsmittel Dichlormethan
5 oder überkritisches Kohlendioxid eingesetzt, wobei beim Einsatz von überkritischem Kohlendioxid die darin enthaltenen Carotinoide anschließend in Dichlormethan überführt werden können oder das Wertprodukt direkt durch Entspannung des Kohlendioxids gewonnen werden kann. Dabei werden die Mengen der Lösungsmittel und Durchmischungszeiten
10 derart gewählt, dass die gewünschte Menge Carotinoide aus der Biomasse extrahiert wird. Insbesondere wird der Extraktionsschritt nur einmal durchgeführt, was technisch und wirtschaftlich sinnvoll ist (vgl. oben).

Zur Durchführung der Extraktion können alle üblichen Verfahren und Apparaturen eingesetzt werden. Insbesondere wird bei nicht getrockneter,
15 aber aufgeschlossener Biomasse eine flüssig/flüssig (Carotinoid liegt in flüssigen Zellbestandteilen gelöst vor und wird daraus extrahiert und bei getrockneter Biomasse eine fest/flüssig Extraktion durchgeführt. Es können Kalt- und Heißextraktion in bestimmten Temperaturbereichen, sowohl
20 kontinuierliche (z.B. Soxhlet-Extraktion, Perforation und Perkolation) als auch diskontinuierliche Verfahren, zu denen beispielsweise Ausschütteln, Auslaugen, Auskochen und Digerieren gehören, verwendet werden. Sie können auch im Gegenstromverfahren durchgeführt werden.

25 Für die flüssig/flüssig Extraktion können beispielsweise Blasensäulen, , pulsierende Kolonnen, Kolonnen mit rotierenden Einbauten, Mixer-Settler-Batterien oder Rührkessel usw. verwendet werden.

Die fest/flüssig Extraktion kann mittels üblicher Apparaturen durchgeführt
30 werden. Vorzugsweise werden Rührkessel oder Mixer-Settler-Apparate eingesetzt.

Alternativ kann der Zellaufschluß ohne vorherige Abtrennung des Fermentationsmediums erfolgen und sich dann eine direkte Trennung einer sich bildenden Carotinoidsuspension von der Biomasse z. B. mittels eines Dekanters durchgeführt werden. Anschließend wird die Carotinoidsuspension in Dichlormethan aufgenommen und weiterverarbeitet oder alternativ durch Wäschen mit verschiedenen wässrigen Lösungen aufgereinigt.

Zur Isolierung der hochreinen Carotinoide aus dem Lösungsmittel wird eine Kristallisation der Carotinoide aus dem verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration durchgeführt. Die verbleibende Mutterlauge kann nach Destillation dem Verfahren erneut zugeführt werden, so dass Produktverluste trotz geringem Aufwand minimiert werden.

Die Kristallisation kann wie üblich erfolgen. Ergänzend wird auf vgl. Römpf Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Kristallisation" verwiesen.

Bevorzugterweise erfolgt die Kristallisation durch graduellen Lösungsmittelaustausch gegen ein Carotinoide nicht lösendes Lösungsmittel. Es wird also kontinuierlich die Löslichkeit der Carotinoide erniedrigt, bis diese als reine Kristalle ausfallen. Hierbei wird vorzugsweise ein "niederer Alkohol" oder Wasser verwendet. Als niederer Alkohol werden aliphatische Alkohole mit 1 bis 4 Kohlenstoffatomen angesehen. Hierzu gehören Methanol, Ethanol, Propanol, Isopropanol, 1-Butanol, tert.-Butanol und sec.-Butanol. Bevorzugterweise wird Methanol eingesetzt.

Die Carotinoid-Lösung kann dabei erwärmt werden, wobei die Temperatur vorzugsweise < 100 °C, insbesondere < 60 °C gehalten wird, so dass Dichlormethan abdestilliert wird. Auch der Einsatz von Vakuum ist denk-

bar. Anschließend werden die Carotinoidkristalle isoliert, welches durch übliche Maßnahmen, insbesondere durch Filtration erfolgen kann. Es können sich, falls gewünscht, weitere optionale Trocknungs- und/oder Reinigungsschritte anschließen. Notwendig sind diese jedoch nicht, da die Carotinoid Kristalle bereits hochrein sind.

Die Carotinoide fallen als hochreine Kristalle an und weisen eine Reinheit von mindestens 95%, bevorzugt > 95%, vorzugsweise > 96%, besonders bevorzugt > 97%, ganz besonders bevorzugt > 98%, höchst bevorzugt > 99% auf.

Die erzielbaren Ausbeuten liegen zwischen 45% und 95%, bevorzugt zwischen 70% und 95% bezogen auf die in der Kulturbrühe vorliegende Menge (0,5 – 15 g/L, bevorzugt 1 – 10 g/L).

Zur Weiterverarbeitung der ebenfalls Carotinoid-haltigen Biomasse zu einem hochwertigen Nahrungsmittel wird zunächst eine Entfernung von Lösemittelresten aus der Carotinoid-haltigen Biomasse vorgenommen. Hierzu erfolgt bevorzugterweise eine Wasserdampfdestillationen bzw. ein so genanntes Strippen mit Wasserdampf (vgl. Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Strippen").

Danach kann gegebenenfalls die Biomasse in der oben abgetrennten Kulturbrühe homogen suspendiert werden, wobei ein Feststoffgehalt > 100 g/L und < 600 g/L eingehalten werden sollte, so daß die nachfolgende Trocknung der Biomasse bzw. Suspension zur Herstellung des Nahrungsmittels ohne technische Schwierigkeiten erfolgen kann. D.h. die Suspension muß pumpbar sein. Als Trocknungsverfahren kommen alle bereits genannten Verfahren und Apparaturen in Frage. Insbesondere wird zur Trocknung die Sprühtrocknung eingesetzt. Dabei kann wie aus der DE 101 04 494 A1 bekannt verfahren werden.

Bei der Sprühtrocknung werden Eingangstemperaturen von ca. 100°C – 180°C, bevorzugt 120°C – 130°C, und Ausgangstemperaturen von ca. 50 – 80°C, bevorzugt 55°C – 70°C gewählt. Als Trocknungsgas wird vorzugsweise Stickstoff eingesetzt.

Das so hergestellte Nahrungsmittel kann entweder direkt verwendet werden oder mittels weiterer Zusätze aufbereitet werden, so wie dies ebenfalls aus der DE 101 04 494 A1 bekannt ist.

10

Als Nahrungsmittel werden Zusammensetzungen angesehen, die der Ernährung dienen. Darunter fallen auch Zusammensetzungen für die Ergänzung der Ernährung. Insbesondere werden als Nahrungsmittel Tierfuttermittel und Tierfutterergänzungsmittel angesehen. Ergänzend wird auf Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Nahrungsmittel" verwiesen.

Das trockene Produkt weist bevorzugt eine Restfeuchte von weniger als 5 % bezogen auf die Trockenmasse auf. Sein Carotinoidgehalt liegt zwischen 0,05 und 20 %, insbesondere 1 und 10 % bezogen auf die Trockenmasse. Der gewünschte Carotinoidgehalt ist über das Ausmaß der Extraktion steuerbar (vgl. oben).

Nach dem erfindungsgemäßen Verfahren erhältliche Nahrungsmittel enthalten also bereits nach der Herstellung große Mengen an Carotinoiden, die nicht zugesetzt werden müssen. Dadurch, dass das Nahrungsmittel neben dem mindestens einen Carotinoid auch Biomasse enthält ist sein Nährstoffgehalt zudem gesteigert. Insbesondere ist der Nährstoffgehalt nach der bevorzugten Alternative stark gesteigert, in dem es neben dem mindestens einen Carotinoid und Biomasse zusätzlich alle Medienbestandteile der Fermentation enthält. Es fallen demnach praktisch keine

Abfälle an, abgesehen von wässrigen Medien, welche jedoch unproblematisch in einer Kläranlage gereinigt werden können. Zusätzlich wird die gesamte Produktionsmenge Carotinoide ohne oder mit nur marginalen Verlusten verwertet, da keine verlustreichen Trenn- bzw. Aufarbeitungsschritte vorgenommen werden müssen, um die gesamte Menge Carotinoid zu extrahieren.

Die in dem oben beschriebenen, erfindungsgemäßen Verfahren eingesetzten Lösungsmittel werden alle soweit wie möglich aufbereitet und anschließend wieder verwendet bzw. dem Verfahren erneut zugeführt. Insbesondere wird das eingesetzte Dichlormethan bereits beim Lösungsmittelaustausch gereinigt und steht anschließend zur erneuten Verwendung bereit. Der niedere Alkohol bzw. das Methanol wird z. B. destillativ gereinigt und ebenfalls wieder verwendet. Als Abfall fallen lediglich der Destillationssumpf an, der zusammen mit den wässrigen Medien gefahrlos einer Kläranlage zugeführt werden kann, wo letztendlich nur eine geringe Menge Klärschlamm als tatsächlicher Abfall anfällt. Somit ist das beschriebene Verfahren im wesentlichen abfallfrei.

20

Die Erfindung wird nachfolgend an Hand von Beispielen näher ausgeführt.

25 **A) Kultivierung von Blakeslea trispora**

Folgende Medien wurden zur Fermentation von Blakeslea trispora zur Produktion der Carotinoide eingesetzt:

Medium 1:

	Glucose	10,00 g/l
30	Baumwollsaatoel	30,00 g/l
	Sojaoel	30,00 g/l

	Dextrin	60,00 g/l
	Baumwollsamemehl	75,00 g/l
	Triton X 100	1,20 g/l
	Ascorbinsäure	6,00 g/l
5	Milchsäure	2,00 g/l
	KH_2PO_4	0,50 g/l
	$\text{MnSO}_4 \times \text{H}_2\text{O}$	100 mg/l
	Thiamin-HCl	2 mg/l
	Isoniazid (Isonicotinsäurehydrazid)	0,75 g/l
10	Der pH wurde auf 6,5 eingestellt.	

Medium 2:

	Glucose	20 g/l
	Asparagin	2,00 g/l
15	KH_2PO_4	5,00 g/l
	$\text{MgSO}_4 \times 7 \text{H}_2\text{O}$	0,50 g/l
	CaCl_2	28 mg/l
	Thiamin-HCl	1,00 mg/l
	Citronensäure	2,00 mg/l
20	$\text{Fe}(\text{NO}_3)_3 \times 9 \text{H}_2\text{O}$	1,50 mg/l
	$\text{ZnSO}_4 \times 7 \text{H}_2\text{O}$	1,00 mg/l
	$\text{MnSO}_4 \times \text{H}_2\text{O}$	0,30 mg/l
	$\text{CuSO}_4 \times 5 \text{H}_2\text{O}$	0,05 mg/l
	$\text{Na}_2\text{MoO}_4 \times 2 \text{H}_2\text{O}$	0,05 mg/l

25

Medium 3

	Glucose	70,00 g/l
30	Asparagin	2,00 g/l
	Hefe Extrakt	1,00 g/l

KH ₂ PO ₄	1,50 g/l
MgSO ₄ x 7 H ₂ O	0,50 g/l
Span 20	1,00 g/l
Thiamin-HCl	5,0 mg/l

- 5 Der pH wurde auf 5,5 eingestellt.

Mit Sporensuspensionen von *Blakeslea trispora* ATCC 14272 Mating Type (-) die 10⁸ (für Medium 2) bzw. 10⁷ (für Medium 1 und 3) Sporen enthielten, wurden je 200 ml der beschriebenen Medien angeimpft. Die Kultivierung erfolgte jeweils in 1-l-Erlenmeyerkolben mit Schikanen. Mit jedem Medium wurden sechs identische Kolben angesetzt und über 7 Tage bei 28°C und 140 UpM im Schüttler inkubiert.

B) Gentechnische Veränderung von *Blakeslea Trispora*

15

Material und Methoden

Molekulargenetische Arbeiten wurden, wenn nicht anders beschrieben, nach den Methoden in Current Protocols in Molecular Biology (Ausubel et al., 1999, John Wiley & Sons) durchgeführt.

20

Stämme und Wachstumsbedingungen

Die *Blakeslea trispora* Stämme ATCC 14271 (Paarungstyp(+)) und ATCC14272 (-) (ein Wildtyp) wurden erhalten. Paarungstyp (-) wurden von der American Type Culture Collection. erhalten. Die Anzucht von *B. trispora* erfolgte in MEP-Medium (Malzextrakt-Pepton-Medium): 30 g/l Malzextrakt (Difco), 3 g/l Pepton (Soytone, Difco), 20 g/l Agar, Einstellung pH 5,5, ad 1000 ml mit H₂O bei 28 °C.

Die Anzucht von *Agrobacterium tumefaciens* LBA4404 erfolgte nach Hoekema et al. (1983, Nature 303:179-180) bei 28 °C für 24. h in Agrobacterien-Minimal Medium (AMM): 10 mM K₂HPO₄, 10 mM KH₂PO₄, 10 mM Glu-

cose, MM-Salze (2,5 mM NaCl, 2 mM MgSO₄, 700 µM CaCl₂, 9 µM Fe-SO₄, 4 mM (NH₄)₂SO₄).

Transformation von *Agrobacterium tumefaciens*

- 5 Das Plasmid pBinAHyg wurde in den Agrobakterienstamm LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der Agrobakterienanzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das *A. tumefaciens* Chromosom), Streptomycin 30 mg/l (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

Transformation von *Blakeslea trispora*

- 15 Zur Transformation wurden die Agrobakterien nach 24 h Anzucht in AMM auf eine OD₆₀₀ von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES (pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 µM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD₆₀₀ von ca. 0,6 angezogen.

- 20 Zur Co-Inkubation von *Blakeslea* ATCC 14271 bzw. ATCC14272 und *Agrobacterium* wurden 100 µl Agrobakteriensuspension mit 100 µl *Blakeslea* Sporensuspension (10⁷ Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt. Zur Selektion auf transformierte *Blakesleazellen* enthielt das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische Selektionsplatten. Gebildete Sporen wurden mit 30 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-

Asparagin, 50 mg/l $\text{MgSO}_4 \times 7\text{H}_2\text{O}$, 150 mg/l KH_2PO_4 , 25 µg/l ThiaminHCl, 100 mg/l Yeast Extract, 100 mg/l Na-desoxycholat, 100 mg/L Hygromycin, 100 mg/L Cefotaxim, pH 5,5, 18 g/l Agar) ausplattiert. Zur Isolierung einzelner gentechnisch veränderter Sporen wurden die Sporen durch ein
5 FACS Gerät der Fa. BectonDickson (Modell Vantage+Diva Option) einzeln auf Selektivmedium abgelegt.

Mutagenese mit MNNG

Zur Reduzierung der Anzahl von Kernen pro Spore wurde eine Behandlung von Sporensuspensionen mit MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) durchgeführt. Hierfür wurde zunächst eine Sporensuspension mit 1×10^7 Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 µg/ml zugegeben. Die Zeit der Inkubation in MNNG wurde so gewählt, dass die
10 Überlebensrate der Sporen ca. 5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und plattiert.
15

Selektion homonukleater Zellen

20 Die Selektion homonukleater Zellen von *Blakeslea trispora* carB^- erfolgte analog zum Versuchsprotokoll für *Phycomyces blakesleeanus* (Roncero et al., 1984, Mutation Research, 125:195-204), modifiziert durch Wachstum in Gegenwart von 5-Carbon-5-Deazariboflavin (1 µg/ml) und Hygromycin 100 (µg/ml).

25

Herstellung genetisch veränderter *Blakeslea trispora* durch Agrobacterium-vermittelte Transformation

Herstellung des rekombinanten Plasmids pBinAHyg

Aus dem Plasmid pANsCos1 (Fig.1, Osiewacz, 1994, Curr. Genet. 26:87-90, SEQ ID NO: 4) wurde die gpdA-hph-trpC-Kassette als BglII/HindIII Fragment isoliert und in das mit BamHI/HindIII geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so erhaltene Vektor wurde als pBinAHyg bezeichnet (Fig. 2, SEQ ID NO: 3) und enthielt das *E. coli* Hygromycin-Resistenzgen (hph) unter Kontrolle des gpd Promotors (SEQ ID NO: 1) und des trpC Terminators (SEQ ID NO: 2) aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium* notwendig sind. Die in den weiter unten beschriebenen Ausführungsbeispielen genannten Vektoren sind Abkömmlinge von pBinAHyg.

Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in *Agrobacterium tumefaciens*

Nachfolgend wird beispielhaft die Übertragung des Plasmids pBinAHyg in Agrobakterien beschrieben. Die Übertragung der Abkömmlinge erfolgte analog.

Das Plasmid pBinAHyg wurde in den Agrobakterienstamm LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der Agrobakterienanzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das *A. tumefaciens* Chromosom), Streptomycin 30 mg/l (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in *Blakeslea trispora*

Zur Transformation wurden die Agrobakterien nach 24 h Anzucht in AMM auf eine OD₆₆₀ von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES

(pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 µM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD₆₆₀ von ca. 0,6 angezogen.

- 5 Zur Co-Inkubation von *Blakeslea trispora* (B.t.) und *Agrobacterium tumefaciens* (A.t.) wurden 100 µl Agrobakteriensuspension mit 100 µl *Blakeslea* Sporensuspension (10^7 Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt.

- Zur Selektion auf transformierte *Blakeslea*-Zellen enthielt das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen
15 Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische Selektionsplatten. Gebildete Sporen wurden mit 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-Asparagin, 50 mg/l MgSO₄ x 7H₂O, 150 mg/l KH₂PO₄, 25 µg/l Thiamin-HCl, 100 mg/l Yeast Extract,
20 100 mg/l Na-desoxycholat, pH 5,5, 100 mg/l Cefotaxim, 100 mg/l Hygromycin, 18 g/l Agar) ausplattiert. Die Übertragung von Sporen auf frische Selektionsplatten wurde dreimal wiederholt. Auf diese Weise wurde die Transformante *Blakeslea trispora* GVO 3005 isoliert. Alternativ erfolgte zur Selektion der GVO (gentechnisch veränderten Organismen) die Einzelablage der Sporen durch den BectonDickinson FACS Vantage+Diva Option
25 auf CM-17 Agar mit 100 mg/l Cefotaxim, 100 mg/l Hygromycin. In diesem Fall wurde nur dort Pilzmycel gebildet, wo die Sporen gentechnisch verändert waren.

- 30 **Nachweis der genetischen Veränderung durch Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in *Blakeslea trispora***

Nachfolgend wird beispielhaft der Nachweis der Übertragung für pBinA-Hyg in *Blakeslea trispora* beschrieben. Der Nachweis der Übertragung der Abkömmlinge erfolgte analog.

- 5 200 ml MEP-Medium (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5) wurden mit 10^5 bis 10^7 Sporen der Transformante *Blakeslea trispora* GVO 3005 beimpft und 7 Tage bei 26 °C mit 200 Upm auf einem Rundschüttler inkubiert. Zum Nachweis der erfolgreichen Transformation wurde DNA aus dem Mycel isoliert (Peqlab Fungal DNA Mini Kit) und in einer PCR (Pro-
- 10 gramm: 94 °C 1 min, dann 30 Zyklen mit 1 min. 94°C, 1 min. 58 °C, 1 min. 72 °C) eingesetzt.

- Zum Nachweis des Hygromycinresistenzgens (hph) wurden die Primer hph-forward (5'-CGATGTAGGAGGGCGTGGATA, SEQ ID NO: 5) und
- 15 hph-reverse (5'-GCTTCTGCGGGCGATTTGTGT, SEQ ID NO: 6) verwendet. Das erwartete Fragment von hph wies eine Länge von 800 bp auf.

- Zur Amplifikation des Kanamycinresistenzgens nptIII und damit als Kontrolle auf Agrobakterien wurden die Primer nptIII-forward (5'-
- 20 TGAGAATATCACCGGAATTG, SEQ ID NO: 7) und nptIII-reverse (5'-AGCTCGACATACTGTTCTTCC, SEQ ID NO: 8) verwendet. Das erwartete Fragment von nptIII wies eine Länge von 700 bp auf.

- Zur Amplifikation eines Fragmentes des Glycerinaldehyd-3-phosphatdehydrogenasegens gpd1 und damit als Kontrolle auf *Blakeslea trispora* wurden die Primer MAT292 (5'-
- 25 GTGAATGGAAATCCCATCGCTGTC, SEQ ID NO: 9) und MAT293 (5'-AGTGGGTACTCTAAAGGCCATACC, SEQ ID NO: 10) verwendet. Das erwartete Fragment von gpd1 wies eine Länge von 500 bp auf.

Das Ergebnis der PCR der *Blakeslea trispora* DNA ist in Fig. 3 anhand eines Standard-Gels gezeigt. Die Spuren des Gels wurden folgendermaßen belegt:

- | | | |
|----|--|--------------------------------|
| 5 | 1) 100 bp Größenmarker (100 bp - 1 kb) | |
| | 2) B.t. GVO 3005 | primer nptIII-for / nptIII-rev |
| | 3) B.t. GVO 3005 | primer hph-for / hph-rev |
| | 4) B.t. GVO 3005 | primer MAT292 / MAT293 (gpd) |
| | 5) A.t. mit Plasmid pBinAHyg | primer nptIII-for / nptIII-rev |
| 10 | 6) A.t. mit Plasmid pBinAHyg | primer hph-for / hph-rev |
| | 7) B.t. 14272 WT | primer nptIII-for / nptIII-rev |
| | 8) B.t. 14272 WT | primer hph-for / hph-rev |
| | 9) B.t. 14272 WT | primer MAT292 / MAT293 (gpd) |
- 15 In der DNA von *Blakeslea trispora* wurde das Hygromycinresistenzgens (hph) und als Positivkontrolle Glycerinaldehyd-3-phosphatdehydrogenasegen (gpd1) nachgewiesen. nptIII konnte demgegenüber nicht nachgewiesen werden.
- 20 Somit wurde die genetische Veränderung von *Blakeslea trispora* durch *Agrobacterium*-vermittelte Transformation nachgewiesen.

Isolierung homokaryotischer GVO von *Blakeslea trispora*:

25 Herstellung homonukleater Stämme

Durch erfolgreichen Transfer des Vectors pBinAHyg und Abkömmlingen von pBinAHyg in *Blakeslea trispora* entstanden entstehen genetisch veränderte Organismen. In GVO von *Blakeslea trispora*. Jedoch liegen in *Blakeslea* liegen in allen Stadien des vegetativen und des sexuellen Zellzyklus mehrkernige Zellen vor. Daher erfolgte erfolgt die Insertion der VectorFremd-DNA in der Regel nur in einem Kern. Ziel ist es aber, dass,

30

Stämme von *Blakeslea* zu erhalten, bei denen die Insertion der Vektor-Fremd-DNA in allen Kernen vorliegt., d.h. Ziel ist ein homonukleates rekombinantes Pilzmycel.

- 5 Zur Herstellung solcher homokaryotischer Zellen wurden zunächst Sporensuspensionen der rekombinanten Stämme mit MNNG behandelt. Hierfür wurde eine Sporensuspension mit 1×10^7 Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 µg/ml zugegeben. Die Dauer der Inkubation mit MNNG
- 10 wurde so gewählt, dass die Überlebensrate der Sporen ~5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und plattiert.

1) Herstellung homonukleater rekombinanter Stämme durch FACS

15 **(fluorescence-activated cell sorting)**

- Ein geringer Anteil der Sporen von *Blakeslea trispora* bzw. der gentechnisch veränderten Stämme von *Blakeslea trispora* ist von Natur aus einkernig. Zur Herstellung homonukleater rekombinanter Stämme, die Fremd-DNA von pBinAHyg oder pBinAHyg-Abkömmlingen enthielten, wurden die
- 20 einkernigen Sporen durch FACS aussortiert und auf MEP (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin plattiert. Die hier gebildeten Mycelien waren homonukleat. Zur Sortierung mit FACS wurden die Sporen eines 3 Tage alten Ausstriches mit 10 ml Tris-HCl 50mMol + 0,1% Span20 pro Agar-Platte abgeschwemmt. Die Sporenkonzentration betrug 0,5 bis $0,8 \times 10^7$ Sporen pro
- 25 ml. Zu 9 ml Sporensuspension wurden 1ml DMSO und 10 µl Syto 11 (Farbstoff-Stammlösung in DMSO Molecular Probes Nr.S-7573) zugegeben. Danach wurde 2 h bei 30°C gefärbt. Die Selektion und Ablage erfolgte mittels eines BectonDickinson FacsVantage+Diva Option. Die Selektion
- 30 erfolgt zuerst nach Größe, um einzelne Sporen von Aggregaten und Verunreinigungen zu trennen. Dann wurden diese Sporen nach ihrer Fluores-

zenz (Anregung = 488nm Emission = 530 nm) sortiert abgelegt. Die linke Schulter der Gauß-Kurve der Fluoreszenzhäufigkeitsverteilung enthielt die einkernigen Sporen.

- 5 Anschließend wurden die Sporen auf MEP-Agarplatten ausplattiert und neue Sporen erzeugt.

Diese Sporen wurden analog zur Vorschrift von Roncero et al. auf Medium mit 5-Carbon-5-deazariboflavin plattiert, das zusätzlich Hygromycin ent-
10 hielt.

Hierdurch wurden homokaryonte Zellen des Genotyps hyg^R und dar^- selektiert.

15

2) Herstellung homonukleater Stämme durch Kernreduktion und Selektion mit FACS

Zur Reduzierung der Anzahl von Kernen pro Spore wurde vor der Selektion eine Behandlung von Sporensuspensionen mit MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) durchgeführt, und so durch chemische Mutagenese eine Kernreduktion erzielt.
20

Hierfür wurde zunächst eine Sporensuspension mit 1×10^7 Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG
25 in einer Endkonzentration von 100 $\mu\text{g/ml}$ zugegeben. Die Zeit der Inkubation in MNNG wurde so gewählt, dass die Überlebensrate der Sporen ca. 5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und nach der unter
30 1) beschriebenen Methode sortiert bzw. selektiert.

30

Alternativ konnten zur Reduktion der Kernzahl in den Sporen auch Röntgen – und UV-Strahlen eingesetzt werden, wie es von Cerdá-Olmedo und Patricia Reau in Mutation Res., 9 (1970), 369-384 beschrieben wurde.

5

3) Herstellung homonukleater Stämme durch Selektion auf rezessive Selektionsmarker

Als rezessiver Selektionsmarker zur Selektion homonukleater Mycelien kommt beispielsweise der rezessive Selektionsmarker pyrG in Frage. Wildtyp-Stämme von *Blakeslea trispora* sind pyrG⁺. Diese Stämme können nicht in Gegenwart des Pyrimidin-Analogs 5-Fluororotat (FOA) wachsen, weil sie FOA durch die Orotidin-5'-monophosphatdecarboxylase zu lethalen Metaboliten umsetzen. Gentechnisch veränderte *Blakesleaa*, die homonukleat pyrG⁻ sind, fehlt die Enzymaktivität Orotidin-5'-monophosphatdecarboxylase. Folglich können diese pyrG⁻-Stämme 5-Fluororotat nicht verwerten. Die Stämme wachsen daher in Gegenwart von FOA und Uracil. Im Fall der Kopplung der Mutation pyrG⁻ und der Insertion von Fremd-DNA auf dem Kern einer einkernigen Spore, kann aus dieser Spore homonukleates rekombinantes Pilzmycel gebildet werden.

20

Zunächst wurde durch Insertion eines Fragments von pyrG (SEQ ID NO: 65) aus *Blakeslea trispora* in pBinAHyg das Plasmid pBinAHygBTpyrG-SCO (SEQ ID NO: 36, Fig. 4) erzeugt. Dieses Plasmid wurde in *Blakeslea trispora* transformiert und führte dort durch homologe Rekombination zur Disruption von pyrG.

25

Homonukleate GVO von *Blakeslea trispora* mit dem Phänotyp pyrG⁻ wurden folgendermaßen selektiert. Zur Agrobakterium-vermittelten Transformation von pBinAHygBTpyrG-SCO wurde wie oben beschrieben auf MEP (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin plattiert. Die Sporen der Transformanten

30

wurden mit 10 ml Tris-HCl 50mM + 0,1% Span20 pro Agar-Platte abgeschwemmt. Die Sporenkonzentration betrug 0,5 bis $0,8 \times 10^7$ Sporen pro ml. Die Sporen wurden anschließend auf FOA-Medium mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin ausplattiert. FOA-Medium enthielt pro Liter 20 g Glucose, 1 g FOA, 50 mg Uracil, 200 ml Citrat-Puffer (0,5 M, pH 4,5) und 40 ml Spurensalzlösung nach Sutter, 1975, PNAS, 72:127). Homonukleate pyrG⁻-Mutanten zeigten Wachstum auf dem Uracil-haltigen FOA-Medium; aber kein Wachstum bei Plattierung auf FOA-Medium ohne Uracil. Auf die gleiche Weise wurden aus den im folgenden beschriebenen GVO von *Blakeslea trispora* zur Herstellung von Xanthophyllen homonukleate GVO hergestellt.

Alternativ ist es möglich die Sporen analog zur Vorschrift von Roncero et al. auf Medium mit 5-Carbon-5-deazariboflavin zu plattieren, das zusätzlich Hygromycin enthält (Roncero et al., 1984, Mutation Research, 125: 195 - 204). Hierdurch werden homokaryonte Zellen des Genotyps hyg^R und dar⁻ selektiert. Nach diesem Prinzip werden homokaryonte Stämme von *Blakeslea trispora* mit dem Phänotyp hyg^R und dar⁻ erzeugt.

20

Ausführungsbeispiele zur Herstellung von gentechnisch veränderten Organismen von *Blakeslea trispora* für die Herstellung von Carotinoiden und Carotinoidvorstufen

Die Erzeugung der im folgenden genannten Plasmide erfolgte durch die Methode „overlap-extension PCR“ und durch anschließende Insertion der Amplifikationsprodukte in das Plasmid pBinAHyg. Die Methode „overlap-extension PCR“ erfolgte wie in Innis et al. (Eds.) PCR protocols: a guide to methods and applications, Academic Press, San Diego beschrieben. Die Transformation der pBinAHyg-Abkömmlinge und die Herstellung homonukleater gentechnisch veränderter Stämme von *Blakeslea trispora* erfolgte wie oben beschrieben.

Gentechnisch veränderte Stämme von *Blakeslea trispora* zur Herstellung von Zeaxanthin

Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung von Zeaxanthin verwendet, codieren also u.a. Hydroxylasen (crtZ):

- 5 - p-tef1-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ (SEQ ID NO: 70) aus *Haematococcus pluvialis* Flotow NIES-144 (Accession No. AF162276) unter Kontrolle des ptef1 Promotors aus *Blakeslea trispora* (Seq. pBinAHygBTpTEF1-HPcrtZ, SEQ ID NO: 37, Fig. 5);
- 10 - p-carRA-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHyg-BTpcarRA-HPcrtZ, SEQ ID NO: 38, Fig. 6)
- 15 - p-carB-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarB aus *Blakeslea trispora* (Seq. pBinAHygBTpcarB-HPcrtZ, SEQ ID NO: 39, Fig. 7)
- 20 - p-carRA-HPcrtZ-TAG-3'carA-IR, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora*. Stromabwärts des Gens der Hydroxylase ist eine Inverted-Repeat-Struktur lokalisiert, die aus dem 3'-Ende von carA und der stromabwärts von carA gelegenen Region stammt (IR, SEQ ID NO: 74, 'Inverted Repeat 1' ca. 350 bp von carA, dann ca. 200 bp 'Loop' und anschließend ca. 350 bp 'Inverted Repeat 2') (Seq. pBinAHyg-BTpcarRA-HPcrtZ-TAG-3'carA-IR, SEQ ID NO: 40, Fig. 8);
- 25 - p-carRA-HPcrtZ-GCG-3'carA-IR, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora*. Das Gen der
- 30

- Hydroxylase ist mit einer Inverted-Repeat-Struktur fusioniert, die aus dem 3'-Ende von *carA* und der stromabwärts von *carA* gelegenen Region stammt (IR, SEQ ID NO: 74, 'Inverted Repeat 1' ca. 350 bp von *carA*, dann ca. 200 bp 'Loop' und anschließend ca. 350 bp 'Inverted Repeat 2'). Das abgeleitete Fusionsprotein besteht folglich aus der Hydroxylase von *Haematococcus pluvialis* und dem Carboxyterminus von *CarA* aus *Blakeslea trispora* (Seq. pBinAHyg-BTp*carRA*-HP*crtZ*-GCG-3'*carA*-IR, SEQ ID NO: 41, Fig. 9);
- 5
- p-*tef1*-EU*crtZ*, enthaltend Gen der Hydroxylase EU*crtZ* (SEQ ID NO: 71) aus *Erwinia uredova* 20D3 (Accession No. D90087) unter Kontrolle des *ptef1* Promotors (Seq. pBinAHygBTpTEF1-EU*crtZ*, SEQ ID NO: 42, Fig. 10);
 - p-*carRA*-EU*crtZ*, enthaltend Gen der Hydroxylase EU*crtZ* aus *Erwinia uredova* 20D3 unter Kontrolle des Promotors *pcarRA* aus *Blakeslea trispora* (Seq. pBinAHygBTp*carRA*-EU*crtZ*, SEQ ID NO: 43, Fig. 11);
 - p-*carB*-EU*crtZ*, enthaltend Gen der Hydroxylase EU*crtZ* aus *Erwinia uredova* 20D3 unter Kontrolle des Promotors *pcarB* aus *Blakeslea trispora* (Seq. pBinAHygBTp*carB*-EU*crtZ*, SEQ ID NO: 44, Fig. 12);
 - p-*gpdA*-HP*crtZ*-t-*crtZ*, enthaltend Gen der Hydroxylase HP*crtZ* aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des *gpdA* Promotors und des Terminators t-*crtZ*; d.h. des stromabwärts von *crtZ* aus *Haematococcus pluvialis* Flotow NIES-144 gelegenen Sequenzabschnitts (SEQ ID NO: 73) (Seq. pBinAHyg-*gpdA*-HP*crtZ*-t*crtZ*, SEQ ID NO: 45, Fig. 13).
 - p-*gpdA*-BT*carR*-HP*crtZ*-BT*carA*, enthaltend Genfusion aus Genen der Lycopincyclase *carR* aus *Blakeslea trispora*, der Hydroxylase HP*crtZ* aus *Haematococcus pluvialis* Flotow NIES-144 und der
- 10
- 15
- 20
- 25

Phytoensynthase *carA* aus *Blakeslea trispora* unter Kontrolle des *gpdA* Promotors aus *Aspergillus nidulans* (Seq. pBinAHyg-*carR_crtZ_carA*, SEQ ID NO: 46, Fig. 14);

5 **Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Canthaxanthin**

Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung von Canthaxanthin verwendet, codieren also u.a. Ketolasen (*crtW*):

- 10 - p-*tef1*-NPcrtW, enthaltend das Gen der Ketolase NPcrtW (SEQ ID NO: 72) aus *Nostoc punctiforme* PCC73102 (ORF148, Accession No. NZ_AABC01000196) unter Kontrolle des *ptef1* Promotors aus *Blakeslea trispora* (Seq. pBinAHygBTpTEF1-NpucrtW, SEQ ID NO: 47, Fig. 15);
- 15 - p-*carRA*-NPcrtW, enthaltend das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 unter der Kontrolle des Promotors *pcarRA* aus *Blakeslea trispora* (Seq. pBinAHygBTp*carRA*-NpucrtW, SEQ ID NO: 48, Fig. 16);
- 20 - p-*carB*-NPcrtW, enthaltend das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 unter der Kontrolle des Promotors *pcarB* aus *Blakeslea trispora* (Seq. pBinAHygBTp*carB*-NpucrtW, SEQ ID NO: 49, Fig. 17);

25 **Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Astaxanthin**

Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung von Astaxanthin verwendet, codieren also u.a. für Hydroxylasen (*crtZ*) und Ketolasen (*crtW*):

- p-carRA-HPcrtZ-pcarRA-NPcrtW, enthaltend das Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 und das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 (ORF148, Accession No. NZ_AABC01000196) beide jeweils unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-HPcrtZ-BTpcarRA-NpucrtW, SEQ ID NO: 50, Fig. 18);
- p-carRA-EUcrtZ-pcarRA-NPcrtW, enthaltend das Gen der Hydroxylase EUcrtZ aus *Erwinia uredovae* 20D3 (Accession No. D90087) und das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 beide jeweils unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-EUcrtZ-BTpcarRA-NpucrtW, SEQ ID NO: 51, Fig. 19);

15 Klonierung und Sequenzanalyse von Genen und Promotoren, die beispielhaft für die gentechnische Veränderung von *Blakeslea trispora* genutzt werden können.

Nachfolgend werden beispielhaft die Klonierung und Sequenzierung verschiedener Gene und Promotoren aus *Blakeslea trispora* beschrieben.

20

Klonierung und Sequenzanalyse ptef1

Die Klonierung von p-tef aus *Blakeslea trispora* erfolgte auf der Grundlage einer bereits in GenBank veröffentlichten Sequenz des Strukturgens für den Translations-Elongationsfaktor 1- α aus *Blakeslea trispora* (AF157235). Ausgehend von dem Sequenzeintrag AF157235 wurden Primer für die inverse PCR ausgewählt, um die stromaufwärts des Strukturgens gelegene Promotoreregion zu amplifizieren und zu sequenzieren.

In der inversen nested PCR an 200 ng XhoI-gespaltener und zirkularisierter genomischer DNA von *Blakeslea trispora* ATCC14272 wurde ein 3000-bp-Fragment in folgendem Ansatz erhalten: Matrizen-DNA (1 μ g genomi-

30

sche DNA von *Blakeslea trispora* ATCC 14272) Primer MAT344 5'-GGCGTACTTGAAGGAACCCTTACCG-3' (SEQ ID NO: 63) und MAT 345 5'-ATTGATGCTCCCGGTCACCGTGATT-3' (SEQ ID NO: 64) je 0,25 µM, 100 µM dNTP, 10 µl Herculase-Polymerasepuffer 10x, 5 U Herculase (Zu-
5 gabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s. 72 °C, 60 s, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus). Der Sequenzabschnitt, der stromaufwärts des vermutlichen Startcodons des Gens *tef1* innerhalb 3000-bp-Fragmentes liegt, wurde als Promotor *ptef1* bezeichnet.

10

Klonierung Sequenzanalyse des Gens der HMG-CoA-Reduktase aus *Blakeslea trispora*

Zunächst wurde mit dem Cosmidvektor pANsCos1 eine Genbank von *Blakeslea trispora* ATCC 14272, Mating Type (–) hergestellt. Der Vektor wurde durch Spaltung mit XbaI linearisiert und anschließend dephosphoryliert.
15 Eine weitere Spaltung mit BamHI schuf die Insertionsstelle, in welche die mit Sau3AI partiell gespaltene und dephosphorylierte genomische DNA von *Blakeslea trispora* ligiert wurde. Die derart gebildeten Cosmide wurden anschließend *in vitro* verpackt und in *Escherichia coli* übertragen.
20 Auf der Grundlage der bekannten Sequenz eines Fragmentes des HMG-CoA-Reduktase codierenden Gens aus *Blakeslea trispora* (Eur. J. Biochem 220, 403–408 (1994)) wurde eine 315-bp-DNA-Sonde durch folgende PCR hergestellt. Reaktionsansatz: 1 µg genomische DNA von *Blakeslea trispora* ATCC 14272, Primer MAT314 5'-
25 CCGATGGCGACGACGGAAGGTTGTT-3' [SEQ ID NO 79] und MAT315 5'-CATGTTTCATGCCCATTGCATCACCT-3' [SEQ ID NO 80] je 0,25 µM, 100 µM dNTP, 10 µl Herculase-Polymerasepuffer 10x, 5 U Herculase (Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 58 °C, 30 s. 72 °C, 30 s, 95 °C, 30 s (30
30 Zyklen); 72 °C, 10 min (1 Zyklus).

Mit dieser DNA-Sonde wurde die Cosmid-Genbank durchmustert. Es wurde ein Klon identifiziert, dessen Cosmid mit der DNA-Sonde hybridisierte. Die Insertion dieses Cosmids wurde sequenziert. Die DNA-Sequenz enthielt einen Abschnitt, der dem Gen einer HMG-CoA-Reduktase zugeordnet wurde [HMG-CoA-Red.gb].

Klonierung und Sequenzanalyse *carB*

(*carB* = Gen der Phytoendesaturase aus *Blakeslea trispora*)

Aus dem Sequenzvergleich der Peptidsequenzen von Phytoendesaturasen und dem Vergleich der zugehörigen DNA-Sequenzen von *Phycomyces blakesleeanus*, *Cercospora nicotianae*, *Phaffia rhodozyma* und *Neurospora crassa* wurden die degenerierten Primer MAT182 5'-GCNGARGGNATHHTGGTA-3' (SEQ ID 52) und MAT192 5'-TCNGCNAGRAADATRTTGTG-3' (SEQ ID 53) abgeleitet. Die PCR wurde in 100 µl Ansätzen durchgeführt. Diese enthielten 200 ng genomische DNA von *Blakeslea trispora* ATCC14272, 1 µM MAT182, 1 µM MAT192, 100 µM dNTP, 10 µl Pfu-Polymerasepuffer 10x, 2,5 U Pfu-Polymerase (Zugabe bei 85 °C), H₂O ad 100 µl.

Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 40 °C, 30 s, 72 °C, 30 s, 95 °C, 30 s (35 Zyklen); 72 °C, 10 min (1 Zyklus).

Hiermit wurde ein 358-bp-Fragment erhalten, dessen abgeleitete Peptidsequenz Ähnlichkeit zu den Sequenzen der Phytoendesaturasen aufwies. Durch die Methode der inversen PCR (Innis et al. in PCR protocols: a guide to methods and applications. 1990. S. 219-227) wurden nach dem Prinzip des Chromosome-Walking die Genregionen stromaufwärts und stromabwärts des 350-bp-Fragmentes folgendermaßen amplifiziert, kloniert und sequenziert:

(i) ein 1,1-kbp-Fragment durch PCR mit den Primern MAT219 5'-AAGTGACACCGGTTACACGCTTGTCTT-3' (SEQ ID 54) und MAT

- 220 5'-GCTTATCACCATCTGTTACCTCCTTGC-3' (SEQ ID 55) erhalten aus 200 ng EcoRI-gespaltener und zirkularisierter genomischer DNA von *Blakeslea trispora* ATCC14272, 0,25 µM MAT219, 0,25 µM MAT220, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase (Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s. 72 °C, 60 s, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus),
- 5
- (ii) ein 2,9-kbp-Fragment durch PCR mit den Primern MAT219 und MAT220 erhalten aus 200 ng XbaI-gespaltener und zirkularisierter genomischer DNA von *Blakeslea trispora* ATCC14272, 0,25 µM MAT219, 0,25 µM MAT220, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase (Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s, 72 °C, 3 min, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus);
- 10
- 15

Der klonierte Sequenzabschnitt ist schematisch in Fig. 20 (SEQ ID NO 77) dargestellt. Die Sequenzierung erfolgte in Strang- und Gegenstrangrichtung mit den klonierten Fragmenten sowie mit den PCR-Produkten. Die Sequenz des klonierten Sequenzabschnitts ist in Fig. 21 (SEQ ID NO 78) gezeigt.

20

Sequenzvergleiche

Die Nukleotidsequenz von carB und die Peptidsequenz des abgeleiteten Proteins CarB wurden mit den bekannten Sequenzen verwandter Proteine verglichen. Zum Sequenzvergleich wurden die Programme GAP und BESTFIT eingesetzt.

25

30 **CarB - Identische Aminoacylreste nach GAP**
Programmeinstellungen:

Gap Weight: 8

Length Weight: 2

Average Match: 2.912

Average Mismatch: -2.003

- 5 Dabei wurde folgende Werte für die Übereinstimmung der Aminosäuren zu CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 72,491

Phaffia rhodozyma: 50,460

Neurospora crassa: 47,943

- 10 *Cercospora nicotianae*: 47,740

CarB - Identische Aminoacylreste nach BESTFIT

Programmeinstellungen:

Gap Weight: 8

- 15 Length Weight: 2

Average Match: 2.912

Average Mismatch: -2.003

Dabei wurde folgende Werte für die Übereinstimmung der Aminosäuren zu CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

- 20 *Phycomyces blakesleeanus*: 73,380

Phaffia rhodozyma: 53,175

Neurospora crassa: 51,896

Cercospora nicotianae: 50,791

- 25 **carB - Identische Basen nach GAP**

Programmeinstellungen:

Gap Weight: 50

Length Weight: 3

Average Match: 10.000

- 30 Average Mismatch: 0.000

Dabei wurde folgende Werte für die Übereinstimmung der Basen zu CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

	<i>Phycomyces blakesleeanus</i> :	64,853
	<i>Cercospora nicotianae</i> :	50,143
5	<i>Phaffia rhodozyma</i> :	43,179
	<i>Neurospora crassa</i> :	42,130

carB -Identische Basen nach BESTFIT

Programmeinstellungen:

10	Gap Weight:	50
	Length Weight:	3
	Average Match:	10.000
	Average Mismatch:	-9.000

Dabei wurde folgende Werte für die Übereinstimmung der Basen zu CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

15	<i>Phycomyces blakesleeanus</i> :	68,926
	<i>Phaffia rhodozyma</i> :	62,403
	<i>Neurospora crassa</i> :	60,230
	<i>Cercospora nicotianae</i> :	56,884

20

Klonierung zur Expression von carB

Zur Klonierung und Expression von carB aus *Blakeslea trispora* wurden von dem oben beschriebenen klonierten Sequenzabschnitt aus *Blakeslea trispora* in sechs Leserastern die möglichen Proteinsequenzen abgeleitet.

25 Diese Proteinsequenzen wurden mit den Sequenzen der Phytoendesaturasen aus *Phycomyces blakesleeanus*, *Phaffia rhodozyma*, *Neurospora crassa*, *Cercospora nicotianae* verglichen. Auf der Grundlage des Sequenzvergleiches wurden im klonierten Sequenzabschnitt der genomischen DNA von *Blakeslea trispora* drei Exons identifiziert, die zusammen-

30 gefügt eine codierende Region ergeben, deren abgeleitetes Genprodukt über die gesamte Länge 72,7% identische Aminoacylreste mit der Phy-

toendesaturase CarB aus *Phycomyces blakesleeanus* aufweist. Dieser Sequenzabschnitt aus drei möglichen Exons und zwei möglichen Introns wurde daher als Gen *carB* bezeichnet. Zur Überprüfung der vorhergesagten Genstruktur wurde die codierende Sequenz von *carB* aus *Blakeslea trispora* durch PCR mit cDNA von *Blakeslea trispora* als Matrize und mit

5 den Primern Bol1425 5'-AGAGAGGGATCCTTAAATGCGAATATCGTTGC-3' (SEQ ID 56) und Bol1426 5'-AGAGAGGGATCCATGTCTGATCAAAAGAAGCA-3' (SEQ ID 57) erzeugt. Das erhaltene DNA-Fragment wurde sequenziert. Die Lokalisation von Exons und Introns wurde durch Vergleich der cDNA mit der

10 genomischen DNA von *carB* bestätigt. In Fig. 21 ist die codierende Sequenz von *carB* schematisch dargestellt. Zur Expression von *carB* in *Escherichia coli* wurde zunächst die NdeI-Schnittstelle in *carB* durch die Methode overlap extension PCR entfernt sowie am 5'-Ende des Gens eine

15 NdeI-Schnittstelle und am 3'-Ende eine BamHI-Schnittstelle eingefügt. Das erhaltene DNA-Fragment wurde mit dem Vektor pJOE2702 ligiert. Das erhaltene Plasmid wurde als pBT4 bezeichnet und zusammen mit pCAR-AE in *Escherichia coli* XL1-Blue kloniert. Die Expression erfolgte durch Induktion mit Rhamnose. Der Nachweis der Enzymaktivität erfolgte

20 durch Nachweis der Lycopinsynthese via HPLC. Die Klonierungsschritte sind im folgenden beschrieben:

PCR 1.1:

Ca. 0,5 µg cDNA von *Blakeslea trispora*, 0,25 µM MAT350 5'-

25 ACTTTATTGGATCCTTAAATGCGAATATCGTTGCTGC-3' (SEQ ID 58),

0,25 µM MAT244 5'-

GTTCCAATTGGCCACATGAAGAGTAAGACAGGAAACAG-3' (SEQ ID 59), 100 µM dNTP, 10 µl Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100µL.

30 Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 40 °C 30s, 4. 72 °C 1 min 30 s, 5. 95 °C 30 s, 6. 50 °C 30 s, 7. 72 °C 1 min 30 s, 8. 95 °C 30 s, 9. 72 °C 10min
 Zyklen: (1-2.) 1x, (3-5.) 5x, (6-8.) 25x, (9.) 1x

5 PCR1.2:

Ca. 0,5 µg cDNA von *Blakeslea trispora*, 0,25 µM MAT243 5'-CCTGTCTTACTCTTCATGTGGCCAATTGGAACCAACAC-3' (SEQ ID 60), 0,25 µM MAT353 5'-CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3' (SEQ ID 61),
 10 100 µM dNTP, 10 µl Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100 µL.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 40 °C 30s, 4. 72 °C 1 min 30 s, 5. 95 °C 30 s, 6. 50 °C 30 s, 7. 72 °C 1 min 30 s, 8. 95 °C 30s, 9. 72 °C 10min
 15 Zyklen: (1 -2.) 1x, (3-5.) 5x, (6-8.) 25x, (9.) 1x

Reinigung der PCR-Fragmente aus PCR 1.1, 1.2

Dazu wurde PCR 2 zur Herstellung der codierenden Sequenz von *carB* aus *Blakeslea trispora* für die Klonierung in pJOE2702 durchgeführt:

20 Ca. 50 ng Produkt aus PCR 1.1 und ca. 50 ng Produkt aus PCR1.2 mit 0,25 µM MAT350 5'-ACTTTATTGGATCCTTAAATGCGAATATCGTTGCTGC-3' (SEQ ID NO 58), 0,25 µM MAT353 5'-CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3' (SEQ ID NO
 25 61), 100 µM dNTP, 10 µL Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100 µL.

Temperaturprofil:

1. 95°C 10 min, 2. 85 °C 5 min, 3. 59 °C 30 s, 4. 72 °C 2 min, 5. 95 °C 30 s, 6. 72°C 10 min
 30 Zyklen: (1-2.) 1x, (3-5.) 22x, (6.) 1x

Anschließend erfolgte eine Reinigung des erhaltenen Fragmentes (~ 1,7 kbp), eine Ligation in Vektor pPCR-Script-Amp, eine Klonierung in *Escherichia coli* XL1-Blue, Sequenzierung der Insertion, Spaltung mit NdeI und BamHI sowie eine Ligation in pJOE2702. Das erhaltene Plasmid wurde als pBT4 bezeichnet.

Charakterisierung und Nachweis der Enzymaktivität von CarB (Phytoendesaturase)

Das von carB abgeleitete Genprodukt wurde als CarB bezeichnet. CarB weist auf Grundlage der Peptidsequenzanalyse folgende Eigenschaften auf:

	Länge:	582 Aminoacylreste
	Molekulare Masse:	66470
	Isoelektrische Punkt:	6,7
15	Katalytische Aktivität:	Phytoendesaturase
	Edukt:	Phytoen
	Produkt:	Lycopin
	EC-Nummer:	EC 1.14.99-

Der Nachweis der Enzymaktivität erfolgte in vivo. Wenn das Plasmid (pCAR-AE) in *Escherichia coli* XL1-Blue übertragen wird, entsteht der Stamm *Escherichia coli* XL1-Blue (pCAR-AE). Dieser Stamm synthetisiert Phytoen. Wenn zusätzlich das Plasmid pBT4 in *Escherichia coli* XL1-Blue übertragen wird, entsteht der Stamm *Escherichia coli* XL1-Blue (pCAR-AE)(pBT4). Da ausgehend von carB eine enzymatisch aktive Phytoendesaturase gebildet wird, produziert dieser Stamm Lycopin.

Die Plasmide pCAR-AE und pBT4 wurden daher in *Escherichia coli* übertragen. Nach Wachstum in Flüssigkultur wurden die Carotinoide aus den Zellen extrahiert und charakterisiert (vgl. oben).

Durch HPLC Analyse wurde nachgewiesen, daß der Stamm *Escherichia coli* XL1-Blue (pCAR-AE) Phytoen und der Stamm *Escherichia coli* XL1-Blue (pCAR-AE)(pBT4) Lycopin produziert. CarB weist folglich die Enzymaktivität einer Phytoendesaturase auf.

5

Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Phytoen

Nachfolgend werden beispielhaft die Herstellung von gentechnisch veränderten Organismen zur Herstellung von Phytoen beschrieben.

10

Vector pBinAHyg Δ carB zur Erzeugung von carB⁻-Mutanten von *Blakeslea trispora*

Für die Deletion von carB in *Blakeslea trispora* wurde der Vektor pBinAHyg Δ carB (SEQ. ID. NO:62, Fig. 22) konstruiert. Der Vorläufer von pBinAHyg Δ carB ist pBinAHyg (SEQ. ID. NO:3, Fig. 2). pBinAHyg wurde folgendermaßen konstruiert:

Aus dem Plasmid pANsCos1 (SEQ. ID. NO:4, Fig. 1, Osiewacz, 1994, Curr. Genet. 26:87-90) wurde die gpdA-hph Kasette als BglII/HindIII Fragment isoliert und in das BamHI/HindIII geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so erhaltene Vektor wurde als pBinAHyg bezeichnet und enthält das *E. coli* Hygromycin-Resistenzgen (hph) unter Kontrolle des gpd Promotors und des trpC Terminators aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium* notwendig sind.

25

Die Amplifikation der codierenden Sequenz von carB mit den Primern MAT350 (SEQ ID NO 58) und MAT353 (SEQ ID NO 61) mittels PCR wurde mit den folgenden Parametern durchgeführt:

30 50 ng pBT4 mit 0,25 μ M MAT350 5'-ACTTTATTGGATCCTTAAAT-GCGAATATCGTTGCTGC-3', 0,25 μ M MAT353 5'-

CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3', 100 μ M
dNTP, 10 μ L Pfu-Polymerase-Puffer, 2,5 U Pfu-Polymerase (Zugabe bei
85 °C, "hot start") und ad 100 μ L H₂O

Temperaturprofil:

- 5 1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 30s, 4. 72°C 2 min, 5. 95 °C 30s,
6. 72 °C 10 min.

Zyklen: (1.-2.) 1x, (3-5.) 30x, (6.) 1x

- 10 Anschließend erfolgte eine Reinigung des erhaltenen Fragmentes (~ 1,7
kbp), eine Spaltung mit HindIII, eine weitere Reinigung des 364-bp-HindIII-
Fragments-carB, gefolgt von einer Spaltung von pBinAHyg mit HindIII, eine
Ligation von 364-bp-HindIII-Fragments-carB in pBinAHyg, eine Transfor-
mation des Vektors in *Escherichia coli* und eine Isolierung des Konstruktes
und Bezeichnung als pBinAHyg Δ carB wie oben beschrieben. Alternativ
15 erfolgte eine partielle Spaltung mit HindIII und die Klonierung eines größe-
ren HindIII-Fragmentes aus carB in pBinAHyg zur Herstellung von pBinA-
Hyg Δ carB.

20 **Erzeugung von carB⁻-Mutanten von *Blakeslea trispora***

- Zunächst wurde das Plasmid pBinAHyg Δ carB in den Agrobakterienstamm
LBA 4404 übertragen, z. B. durch Elektroporation (vgl. oben). Anschlie-
ßend wurde das Plasmid von *Agrobacterium tumefaciens* LBA 4404 in
Blakeslea trispora ATCC 14272 und in *Blakeslea trispora* ATCC 14271
25 übertragen (vgl. oben). Der erfolgreiche Nachweis des Gentransfers in
Blakeslea trispora erfolgte über Polymerase-Kettenreaktion nach fol-
gendem Protokoll:

- Ca. 0,5 μ g DNA aus *Blakeslea trispora* ATCC 14272 carB⁻ bzw. ATCC
14271 carB⁻ wurden mit 0,25 μ M Primer hph forward 5'-
30 CGATGTAGGAGGGCGTGGATA-3' (SEQ ID NO 5), 0,25 μ M Primer hph
reverse 5'-GCTTCTGCGGGCGATTTGTGT-3' (SEQ ID NO 6), 100 μ M

dNTP, 10 µL Herculanse-Polymerase-Puffer, 2,5 U Herculanse-DNA-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 µl H₂O umgesetzt.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 1 min, 4. 72 °C 1 min, 5. 94 °C 1 min, 6. 72 °C 10 min.

Zyklen: (1.-2.) 1x, (3-5.) 30x, (6.) 1x

Als Negativkontrolle wurde eine Amplifikation des Kanamycinresistenzgens aus *Agrobacterium* versucht. Dazu wurden folgende PCR-

- 10 Bedingungen verwendet:

Ca. 0,5 µg DNA aus *Blakeslea trispora* ATCC 14272 *carB*⁻ bzw. ATCC 14271 *carB*⁻ wurden mit 0,25 µM Primer *nptIII* forward 5'-TGAGAATATCACCGGAATTG-3' (SEQ ID NO 7), 0,25 µM Primer *nptIII* reverse 5'-AGCTCGACATACTGTTCTTCC-3' (SEQ ID NO 8), 100 µM

- 15 dNTP, 10 µL Herculanse-Polymerase-Puffer, 2,5 U Herculanse-DNA-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 µL H₂O umgesetzt.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 1 min, 4. 72 °C 1 min, 5. 94 °C 1 min, 6. 72 °C 10 min-

- 20 Zyklen: (1-2.) 1x, (3-5.) 30x, (6.) 1x

C) Produktion von Carotinoiden und Carotinoidvorstufen mit *Blakeslea trispora*

- 25 Zur Produktion der Carotinoide Zeaxanthin, Canthaxanthin, Astaxanthin und Phytoen wurden die entsprechenden gentechnisch veränderten *Blakeslea trispora* (+) und (-) Stämme fermentiert, das produzierte Carotinoid mittels HPLC Analyse nachgewiesen und isoliert.

Das Flüssigmedium zur Produktion von Carotinoiden enthielt pro Liter: 19 g Maismehl, 44 g Sojamehl, 0,55 g KH_2PO_4 , 0,002 g Thiaminhydrochlorid, 10 % Sonnenblumenöl. Der pH wurde mit KOH auf 7,5 eingestellt.

- 5 Zur Herstellung der Carotinoiden wurden Schüttelkolben mit Sporensuspensionen von (+) und (-) Stämmen der GVO von *Blakeslea trispora* beimpft. Die Schüttelkolben wurden bei 26 °C mit 250 rpm für 7 Tage inkubiert. Alternativ wurde zu Mischungen der Stämme nach 4 Tagen Trisporsäuren zugegeben und weitere 3 Tage inkubiert. Die Endkonzentration der
- 10 Trisporsäuren betrug 300 - 400 µg/ml.

Extraktion und Analytik

Extraktion:

1. Entnahme von 10 ml Kultursuspension
- 15 2. Zentrifugation, 10 min, 5.000 x g
3. Verwerfen des Überstandes
4. Resuspendierung des Pellets in 1 ml Tetrahydrofuran (THF) durch Vortexen
5. Zentrifugation, 5 min, 5.000 x g
- 20 6. Abnahme der THF-Phase
7. Wiederholung der Schritte 4.-6. (2 x)
8. Vereinigung der THF-Phasen
9. Zentrifugation der vereinigten THF-Phasen 5 min bei 20.000 x g, um Reste der wäßrigen Phase abzutrennen

25

Analytik

Messung von Phytoen mittels HPLC

- Säule: ZORBAX Eclipse XDB-C8, 5 µm, 150*4,6 mm
- 30 Temperatur: 40 °C
- Flußrate: 0,5 ml/min

Injektionsvolumen: 10 µl

Detektion: UV 220 nm

Stoppzeit: 12 min

Nachlaufzeit: 0 min

5 Maximaldruck: 350 bar

Eluent A: 50 mM NaH₂PO₄, pH 2,5 mit Perchlorsäure

Eluent B: Acetonitril

Gradient:

	Zeit [min]	A [%]	B [%]	Fluß [ml/min]
10	0	50	50	0,5
	12	50	50	0,5

Als Matrix wurden Extrakte der Fermentationsbrühen verwendet. Vor der HPLC wurde jede Probe durch ein 0,22 µm Filter filtriert. Die Proben wurden kühl gehalten und vor Licht geschützt. Zur Kalibrierung wurden jeweils 50 - 1000 mg/l eingewogen und in THF gelöst. Als Standard wurde Phytoen verwendet, welches unter den gegebenen Bedingungen eine Retentionszeit von 7,7 min. aufweist.

20 **Messung von Lycopin, β-Carotin, Echinenon, Canthaxanthin, Cryptoxanthin, Zeaxanthin und Astaxanthin mittels HPLC**

Säule: Nucleosil 100-7 C18, 250*4,0 mm (Macherey & Nagel)

Temperatur: 25 °C

Flußrate: 1,3 ml/min

25 Injektionsvolumen: 10 µl

Detektion: 450 nm

Stoppzeit: 15min

Nachlaufzeit: 2 min

Maximaldruck: 250 bar

30 Eluent A: 10% Aceton, 90% H₂O

Eluent B: Aceton

Gradient:

	Zeit [min]	A [%]	B [%]	Fluß [ml/min]
	0	30	70	1,3
	10	5	95	1,3
5	12	5	95	1,3
	13	30	70	1,3

Als Matrix wurden Extrakte der Fermentationsbrühen verwendet. Vor der HPLC wurde jede Probe durch ein 0,22 µm Filter filtriert. Die Proben wurden kühl gehalten und vor Licht geschützt. Zur Kalibrierung wurden jeweils 10 mg eingewogen und in 100 ml THF gelöst. Als Standard wurden folgende Carotinoide mit folgenden Retentionszeiten eingesetzt β-Carotin (12,5 min), Lycopin (11,7 min), Echinenon (10,9 min), Cryptoxanthin (10,5 min), Canthaxanthin (8,7 min), Zeaxanthin (7,6 min) und Astaxanthin (6,4 min) [s. Fig. 23].

Produktion von Zeaxanthin mit gentechnisch veränderten Stämmen von *Blakeslea trispora*

Nachfolgend wird beispielhaft die Herstellung von Zeaxanthin mit gentechnisch veränderten Organismen (GVO) von *Blakeslea trispora* beschrieben.

Durch Agrobakterium-vermittelte Transformation wurde der Vektor pBinA-HygBTpTEF1-HPcrtZ in *Blakeslea trispora* übertragen (s.o.). Ein Hygromycin-resistenter Klon wurde isoliert und auf eine Kartoffel-Glucose-Agarplatte (Merck KGaA, Darmstadt) übertragen.

Nach drei Tagen Inkubation bei 26°C wurde ausgehend von dieser Platte eine Sporensuspension hergestellt. Ein 250-ml-Erlenmeyerkolben ohne Schikanen mit 50 ml Growth-Medium (Maismehl 47 g/l, Sojamehl 23 g/l, KH₂PO₄ 0,5 g/l, Thiamin-HCl 2,0 mg/l, pH mit NaOH vor der Sterilisation auf 6,2–6,7 eingestellt) wurde mit 1x10⁵ Sporen beimpft. Diese Vorkultur inkubierte 48 Stunden bei 26 °C und 250 upm. Für die Hauptkultur

wurde ein 250-ml-Erlenmeyerkolben ohne Schikane enthaltend 40 ml Produktionsmedium mit 4 ml der Vorkultur beimpft und 8 Tage bei 26 °C und 150 upm inkubiert. Das Produktionsmedium enthielt Glucose 50 g/l, Casein Acid Hydrolysate 2 g/l, Hefeextrakt 1 g/l, L-Asparagin 2 g/l, KH₂PO₄ 5 g/l, MgSO₄ x 7 H₂O 0,5 g/l, Thiamin-HCl 5 mg/l, Span20 10 g/l, Tween 80 1 g/l, Linolsäure 20 g/l, Maisquellwasser 80 g/l. Nach 72 Stunden erfolgte die Zugabe von Kerosin in einer Endkonzentration von 40 g/l Kerosin.

Nach der Ernte der Kulturen werden die verbliebenen ungefähr 35 ml Kultur mit Wasser auf 40 ml aufgefüllt. Anschließend werden die Zellen im Hochdruckhomogenisator, Typ Micron Lab 40, Fa. APV Gaulin, 3 x bei 1500 bar aufgeschlossen.

Die Suspension mit den aufgeschlossenen Zellen wurde mit 35 ml THF versetzt und 60 min bei RT im Dunkeln bei 250 upm geschüttelt. Danach wurden 2 g NaCl zugegeben und das Gemisch nochmals geschüttelt. Der Extraktionsansatz wurde dann 10 min bei 5000 x g zentrifugiert. Die gefärbte THF-Phase wurde abgenommen, die Zellmasse war vollständig entfärbt.

Die THF-Phase wurde am Rotationsverdampfer bei 30 mbar und 30 °C auf 1 ml eingeeengt und danach nochmals in 1 ml THF aufgenommen. Nach Zentrifugation 5 min bei 20 000 x g wurde ein Aliquot der oberen Phase entnommen und durch HPLC analysiert (Fig. 24, Fig. 23).

D) Aufarbeitung und Isolierung der Carotinoide bzw. des Nahrungsmittels

Die oben unter A) angegebenen Kulturbrühen wurden wie nachfolgend aufgearbeitet, um hochreine Carotinoide und ein entsprechendes Nahrungsmittel zu erhalten.

Der Carotinoidgehalt der Kulturbrühen 1, 2, 3 betrug zwischen 0,5 und 1,5 g/L.

D1) Beispiel gemäß Variante a) IIA und Variante b) IIA bzw. IIB

- 5 Die Kulturen mit identischen Medien (insgesamt ca. 1 L) wurden am Ende des Kultivierungszeitraums vereinigt und mit Hilfe eines Dispergiergeräts (Ultra.Turrax ®) homogenisiert.

Die Feststoffkonzentration in den Medien 1 und 2 betrug 37 g/l bzw. 11
10 g/L. Die Entwässerung der Kulturbrühe erfolgte durch eine Zentrifuge. Bei hohen Zellkonzentrationen bzw. hohem Feststoffgehalt des Mediums kann die Kulturbrühe auch ohne vorherige fest-flüssig-Trennung weiterverarbeitet werden (Medium 3: 127 g Feststoff/L. Nach vorheriger Homogenisation mit einem Dispergiergerät (Ultra-Turrax ®) und unter ständigem Rühren
15 der Suspension wurde die Zellmasse über eine Schlauchpumpe auf den Trockner aufgegeben. Die Eindüsung in den Zylinder des Laborsprühtrockners erfolgte dabei über eine Zweistoffdüse mit dem Durchmesser 2,0 mm. Eingedüst wurde mit 2 bar und 4,5 Nm³/h Stickstoff. Die Temperatur am Eintritt betrug ca. 125°C bis 127°C. Das Trocknungsgas war Stickstoff
20 mit einer Flussrate von 22 Nm³/h. Die Austrittstemperatur betrug zwischen 59°C und 61°C. Bei jeder der drei Fermentationsbrühen konnte am Zyklon des Sprühtrockners rieselfähiges Produkt abgeschieden werden. Die Wandbeläge im Turm (sofern vorhanden) platzten automatisch von der Gefäßwand ab und werden als unproblematisch eingestuft.

25

Es wurden zwischen 8 und 100 g pulveriges Nahrungsmittel erhalten, welches direkt als Tierfuttermittel verwendet werden könnte. Es enthielt ca. 1-10 % Carotinoide bezogen auf das Trockengewicht. Die Restfeuchte betrug weniger als 5%.

30

Beispiel gemäß Variante b) IIC**D2) Extraktion mit Tetrahydrofuran**

Die Zellen aus je 40 ml der Kulturbrühen 1, 2, 3 wurden 3 x bei 1500 bar
5 durch einen Hochdruckhomogenisator, Typ Micron Lab 40, Fa. APV Gaulin aufgeschlossen. Je 20 ml der Suspensionen mit den aufgeschlossenen Zellen wurden mit 20 ml Tetrahydrofuran versetzt und 30 min, bei 30°C im Rundschüttler bei 200 Upm geschüttelt. Danach wurden 2 g NaCl zuge-
10 setzt und zur Phasentrennung 5 min bei 5000 x g zentrifugiert. Die THF-Phase wurde abgenommen. Danach wurde die wässrige Phase nochmals mit 20 ml THF extrahiert. Die Extrakte wurde vereinigt. Die Carotinoidkonzentration wurde durch HPLC quantifiziert.

D3) Extraktion mit Methylenchlorid

15 Die Biomassenabtrennung aus der Kulturbrühe (200 mL) erfolgte durch Zentrifugation bei 5.000 x g für 10 min. in einer Laborzentrifuge.

Die abgetrennte Biofeuchtmasse (jeweils ca. 10 g bis 100 g) wurde mit 10
- 100 mL Wasser vermischt, um wasserlösliche Komponenten zu entfer-
20 nen. Die Biomasse wurde abgetrennt (Laborzentrifuge) und danach mit Dampf (T = 121, t = 30 min, 1 bar) im Autoklaven sterilisiert und so die Zellen aufgeschlossen.

Zu den Zelltrümmern wurden 25 - 250 g Methylenchlorid zugegeben und
25 das Carotinoid aus der Biomasse mittels Ausschütteln extrahiert. Die Biomasse wurde in einer Laborzentrifuge abgetrennt.

Es wurde ein Lösungsmitteltausch von Methylenchlorid zu Methanol durchgeführt, wozu die Carotinoidlösung ca. vier Stunden bei 40°C bis
30 60°C gehalten und über diesen Zeitraum kontinuierlich mit insgesamt 20 - 200 mL Methanol versetzt wurde. Methylenchlorid wurde dabei als Lö-

5 sungsmittel zurückgewonnen. Erste Carotinoid Kristalle fielen aus. Anschließend wurde langsam, über 6 h auf ca. 10 °C abgekühlt, wobei die Carotinoid Kristalle an Größe und Anzahl zunahmen. . Danach wurde die Mutterlauge abfiltriert und die Carotinoid Kristalle getrocknet. Ein Teil der Mutterlauge kann zum Lösungsmitteltausch wiederverwendet werden. Der andere Teil wird destilliert und das so gereinigte Methanol im Lösungsmitteltausch wiederverwendet.

10 Es wurden 0,0,08 g bis 0,24 g Carotinoid Kristalle erhalten, welche eine Reinheit (HPLC, vgl. oben) von 95 % aufwiesen. Die Ausbeute an Carotinoid Kristallen betrug 80 % bezogen auf die Konzentration an Carotinoid in der Biomasse.

15 Die abgetrennte methylenchloridfeuchte Biomasse wurde nach Wasserdampfdestillation sprühgetrocknet ($T_E = 125\text{ °C}$, $T_A = 60\text{ °C}$) und kann als Tierfuttermitteladditiv eingesetzt werden.

20 Hierzu wurde nach vorheriger Homogenisation mit einem Dispergiergerät (Ultra-Turrax) und unter ständigem Rühren der Suspension die Zellmasse über eine Schlauchpumpe auf den Trockner aufgegeben.

25 Die Eindüsung in den Zylinder des Laborsprührockners erfolgte dabei über eine Zweistoffdüse mit dem Durchmesser 2,0 mm. Eingedüst wurde mit 2 bar und 4,5 Nm³/h Stickstoff. Die Temperatur am Eintritt betrug ca. 125°C bis 127°C. Das Trocknungsgas war Stickstoff mit einer Flussrate von 22 Nm³/h. Die Austrittstemperatur betrug zwischen 59°C und 61°C. Bei jeder der drei Fermentationsbrühen konnte am Zyklon des Sprührockners rieselfähiges Produkt abgeschieden werden. Die Wandbeläge im Turm (sofern vorhanden) platzten automatisch von der Gefäßwand ab und wurden als unproblematisch eingestuft.

30

Es wurden ca. 2,5 – 25 g pulveriges Nahrungsmittel erhalten, welches direkt als Tierfuttermittel verwendet werden könnte. Es enthielt ca. 0,5% - 1,5% Carotinoide bezogen auf das Trockengewicht. Die Restfeuchte betrug weniger als 5%.

5

Insgesamt (einschließlich des aufgereinigten Carotinoid-Nahrungsmittels) betrug die Ausbeute an Carotinoid ca. 95 % bezogen auf die Ausgangsmenge Carotinoid in der Kulturbrühe.

10

15

20

25

30

Patentansprüche

1. Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderter Organismen der Gattung *Blakeslea* umfassend
- 5
- (i) Transformation mindestens einer der Zellen,
- (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder in mehreren genetischen Merkmalen alle gleichartig verändert sind und diese
- 10 genetische Veränderung zur Ausprägung bringen, und
- (vi) Selektion und Vermehrung der gentechnisch veränderten Zelle oder Zellen,
- (vii) Kultivierung der gentechnisch veränderten Zellen,
- (viii) Bereitstellung des von den gentechnisch veränderten Zellen
- 15 produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe.
2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** es sich um Zellen von Pilzen der Art *Blakeslea trispora* handelt.
3. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet, dass** in
- 20 der Transformation (i) ein Vector oder freie Nukleinsäuren verwendet werden.
4. Verfahren nach Anspruch 3, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector in das Genom mindestens einer der Zellen integriert wird.

5. Verfahren nach Anspruch 4, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector einen Promotor und/oder einen Terminator enthält.
6. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 5, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector enthaltend den gpd, pcarB, pcarRA und/oder ptef1 Promotor und/oder den trpC Terminator eingesetzt wird.
7. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 6, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector enthaltend ein Resistenzgen eingesetzt wird.
8. Verfahren nach Anspruch 7, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector ein Hygromycin-Resistenzgen (hph), insbesondere aus E. coli enthält.
9. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch gekennzeichnet, dass** der gpd Promotor die Sequenz SEQ ID NO: 1 aufweist.
10. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch gekennzeichnet, dass** der trpC Terminator die Sequenz SEQ ID NO: 2 aufweist.
11. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch gekennzeichnet, dass** der tef1 Promotor die Sequenz SEQ ID NO: 35 aufweist.
12. Verfahren nach einem der Ansprüche 6 bis 11, **dadurch gekennzeichnet, dass** der gpd Promotor und der trpC Terminator aus *Aspergillus nidulans* stammen.

13. Verfahren nach einem Ansprüche 3 bis 12, **dadurch gekennzeichnet, dass** der Vector die SEQ ID NO: 3 umfasst.
14. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Transformation (i) mittels Agrobakterien, Konjugation, Chemikalien, Elektroporation, Beschuss mit DNA-beladenen Partikeln, Protoplasten oder Mikroinjektion durchgeführt wird.
15. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** in der Homokaryontisierung (ii) ein mutagenes Agens eingesetzt wird.
16. Verfahren nach Anspruch 15, **dadurch gekennzeichnet, dass** als mutagenes Agens N-Methyl-N'-nitro-nitrosoguanidin (MNNG), UV-Strahlung oder Röntgenstrahlung eingesetzt wird.
17. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Selektion durch Markierung und/oder Auswahl der einkernigen Zellen erfolgt.
18. Verfahren nach einem der vorhergehenden Ansprüche 1 - 17, **dadurch gekennzeichnet, dass** in der Selektion 5-Carbon-5-deazariboflavin (darf) und Hygromycin (hyg) oder 5-Fluororotat (FOA) und Uracil und Hygromycin eingesetzt werden.
19. Verfahren nach einem der Ansprüche 3 bis 18, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder deren Vorstufen enthält.
20. Verfahren nach einem der Ansprüche 3 bis 19, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinen oder Xanthophyllen enthält.

21. Verfahren nach einem der Ansprüche 3 bis 20, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β -Carotin, α -Carotin, Lutein, Phytofluen, Bixin oder Phytoen enthält.

10

22. Verfahren zur Bereitstellung mindestens eines hochreinen Carotinoids und eines Nahrungsmittels, enthaltend Carotinoide-produzierende Organismen und mindestens das eine Carotinoid, umfassend nach der Kultivierung von Carotinoide-produzierenden Organismen der Gattung Blakeslea die Schritte

- I) Abtrennung der Biomasse,
IA) ggf. Waschen der Biomasse mit einem Carotinoide nicht lösenden Lösungsmittel, insbesondere Wasser,
IB) Sterilisation und Zellaufschluß der Biomasse,
IC) ggf. Trocknung und/oder homogene Verteilung und
II) partielle Extraktion der Carotinoide aus der aufgeschlossenen Biomasse mittels eines Carotinoide lösenden Lösungsmittels und Trennung des Lösungsmittels von der Biomasse,
IIA)

- 1) Entfernung von Lösemittelresten aus der Carotinoid-haltigen Biomasse,
- 2) ggf. homogene Suspension der Biomasse mit einem Biomasse-Feststoffgehalt > 10
- 5 3) Trocknung der Biomasse bzw. Suspension zur Herstellung des Nahrungsmittels,

IIB)

- 1) Kristallisation der Carotinoide aus dem verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration.
- 10

23. Verfahren nach Anspruch 22, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid aus der Gruppe bestehend aus Carotinen und Xanthophyllen ausgewählt ist.

24. Verfahren nach Anspruch 22 oder 23, **dadurch gekennzeichnet, dass**

15 das mindestens eine Carotinoid aus der Gruppe bestehend aus Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β -Carotin, Lutein, Phytofluen, Bixin und Phytoen ausgewählt ist.

20 25. Verfahren nach einem der Ansprüche 22 bis 24, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid Astaxanthin, Zeaxanthin, *Bixin* oder Phytoen ist.

26. Verfahren nach einem der Ansprüche 22-25, **dadurch gekennzeichnet, dass** die Sterilisation und der Zellaufschluß mittels Wasserdampf

25 oder Mikrowellenstrahlung durchgeführt werden.

27. Verfahren nach einem der Ansprüche 22-26, **dadurch gekennzeichnet, dass** die Extraktion der Carotinoide aus der Biomasse mittels Me-

thylenchlorid oder überkritischem Kohlendioxid oder Tetrahydrofuran durchgeführt wird.

28. Verfahren nach Anspruch 27, **dadurch gekennzeichnet, dass** die im überkritischen Kohlendioxid gelösten Carotinoide direkt isoliert werden
5 oder in Methylenchlorid aufgenommen werden.

29. Verfahren nach einem der Ansprüche 22-28, **dadurch gekennzeichnet, dass** die Extraktion der Carotinoide aus der Biomasse ein oder ggf. mehrstufig erfolgt.

30. Verfahren nach einem der Ansprüche 22-29, **dadurch gekennzeichnet, dass** die Entfernung von Lösungsmitteln aus der Biomasse im
10 Schritt IA1) mittels Wasserdampf-Destillation .

31. Verfahren nach einem der Ansprüche 22-30, **dadurch gekennzeichnet, dass** die Trocknung in Schritt IIA3) mittels Sprühtrocknung oder Kontakt Trocknung durchgeführt wird.

15 32. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Kristallisation im Schritt IIB1) durch graduellen Lösungsmittelaustausch gegen ein Carotinoide nicht lösendes Lösungsmittel erfolgt.

20 33. Verfahren nach Anspruch 32, **dadurch gekennzeichnet, dass** der Austausch des verwendeten Lösungsmittels gegen Wasser oder einen niederen Alkohol, insbesondere Methanol erfolgt.

34. Verfahren nach Anspruch 13, **dadurch gekennzeichnet, dass** der gentechnisch veränderte Organismus der Gattung Blakeslea durch Transformation mit einem Vector, der eine Sequenz aus der Gruppe
25 bestehend aus den SEQ ID NO: 37 – 51 und 62 aufweist, herstellbar ist.

35. Verfahren zur Herstellung eines Nahrungsmittels enthaltend Organismen der Gattung *Blakeslea* und mindestens ein Carotinoid, umfassend nach der Kultivierung von Carotinoide-produzierenden Organismen der Gattung *Blakeslea* die Schritte

- 5 I) Homogene Suspendierung der Feststoffe der Kulturbrühe und
- IIA) bei einem Biomasse-Feststoffgehalt der Kulturbrühe von > 2 %
- 10 1) ggf. Konzentration der Kulturbrühe auf einen Feststoffgehalt < 50 % und
- 2) Trocknung der Kulturbrühe zur Herstellung des Nahrungsmittels
- oder
- 15 IIB) bei einem Feststoffgehalt von < 2 % der Kulturbrühe,
- 1) Konzentration der Kulturbrühe auf einen Feststoffgehalt > 2 % und < 50 % und
- 2) Trocknung der Suspension zur Herstellung des Nahrungsmittels,
- 20 oder
- IIC) unabhängig vom Feststoffgehalt der Kulturbrühe,
- 1) Abtrennung der Biomasse,
- 25 2) ggf. Waschen der Biomasse mit Carotinoide nicht lösenden Lösungsmitteln, insbesondere Wasser,
- 3) Sterilisation und Zellaufschluß,
- 4) ggf. Trocknung und homogene Verteilung,
- 5) partielle Extraktion der Carotinoide aus der Biomasse
- 30 mittels eines Carotinoide lösendes Lösungsmittels,

- 5a) Abtrennung der Carotinoid-haltigen Biomasse vom Carotinoid-haltigen Lösungsmittel,
5b) Entfernung von Lösemittelresten aus der Biomasse und
5c) Trocknung der Biomasse zur Herstellung des Nahrungsmittels,
6) Kristallisation der Carotinoide aus dem in 5a) verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration.
- 10 36. Verfahren nach Anspruch 35, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid aus der Gruppe bestehend aus Carotinen und Xanthophyllen ausgewählt ist.
- 15 37. Verfahren nach Anspruch 35 oder 36, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid aus der Gruppe bestehend aus Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Anthonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β -Carotin, Lutein, Bixin, Phytoen ausgewählt ist.
- 20 38. Verfahren nach einem der Ansprüche 35-37, **dadurch gekennzeichnet, dass** das mindestens eine Carotinoid Astaxanthin, Zeaxanthin, Bixin oder Phytoen ist.
39. Verfahren nach einem der Ansprüche 35-38, **dadurch gekennzeichnet, dass** die Sterilisation und der Zellaufschluß im Schritt II3) mittels Wasserdampf oder Mikrowellenstrahlung durchgeführt wird.
- 25 40. Verfahren nach einem der Ansprüche 35-39, **dadurch gekennzeichnet, dass** die Extraktion der Carotinoide aus der Biomasse im Schritt II5) mittels Methylenchlorid oder überkritischen Kohlendioxid durchgeführt wird.

41. Verfahren nach Anspruch 40, **dadurch gekennzeichnet, dass** die im überkritischen Kohlendioxid gelösten Carotinoide direkt isoliert werden oder in Methylenchlorid aufgenommen werden..
- 5 42. Verfahren nach einem der Ansprüche 35-41, **dadurch gekennzeichnet, dass** die Extraktion der Carotinoide aus der Biomasse ein- oder ggf. mehrstufig erfolgt.
43. Verfahren nach einem der Ansprüche 35-42, **dadurch gekennzeichnet, dass** die Entfernung von Lösungsmitteln aus der Biomasse im Schritt IIC5b) mittels Wasserdampf-Destillation.
- 10 44. Verfahren nach einem der Ansprüche 35-43, **dadurch gekennzeichnet, dass** die Trocknung in einem der Schritte IIA1), IIB2) oder IIC5c) mittels Sprühtrocknung oder Kontakt durchgeführt wird.
45. Verfahren nach einem der Ansprüche 35-44, **dadurch gekennzeichnet, dass** die Kristallisation im Schritt IIC6) durch graduellen Lösungsmittelaustausch gegen ein Carotinoide nicht lösendes Lösungsmittel erfolgt.
- 15 46. Verfahren nach Anspruch 45, **dadurch gekennzeichnet, dass** der Austausch des verwendeten Lösungsmittels gegen Wasser oder einen niederen Alkohol, insbesondere Methanol erfolgt.
- 20 47. Verfahren nach einem der Ansprüche 35-46, **dadurch gekennzeichnet, dass** der gentechnisch veränderte Organismus der Gattung *Blakeslea* durch Transformation mit einem Vector, der eine Sequenz aus der Gruppe bestehend aus den SEQ ID NO: 37 – 51 und 62 aufweist, herstellbar ist.
- 25 48. Nahrungsmittel, insbesondere Tierfuttermittel herstellbar nach einem der Verfahren der Ansprüche 1 bis 47.

49. Nahrungsergänzungsmittel, insbesondere Tierfutterergänzungsmittel herstellbar nach einem der Verfahren der Ansprüche 1 bis 47.
50. Verfahren nach einem der Ansprüche 1-49 dadurch gekennzeichnet, daß Nahrungsmittel und Tierfuttermittel aus einer Fermentation erhältlich sind.
51. Verfahren nach einem der Ansprüche 1-49 dadurch gekennzeichnet, daß Nahrungsergänzungsmittel und Tierfutterergänzungsmittel aus einer Fermentation erhältlich sind.
52. Verfahren nach einem der Ansprüche 1-49 dadurch gekennzeichnet, daß mindestens zwei Produkte aus der Gruppe Nahrungsmittel, Nahrungsergänzungsmittel, Tierfuttermittel und Tierfutterergänzungsmittel aus einer Fermentation erhältlich sind.
53. Verwendung der nach einem der Verfahren der Ansprüche 1 bis 14 erhältlichen Carotinoide zur Herstellung von kosmetischen, pharmazeutischen, dermatologischen Zubereitungen, Nahrungsmitteln, Nahrungsergänzungsmitteln, Tierfuttermittel oder Tierfutterergänzungsmitteln.

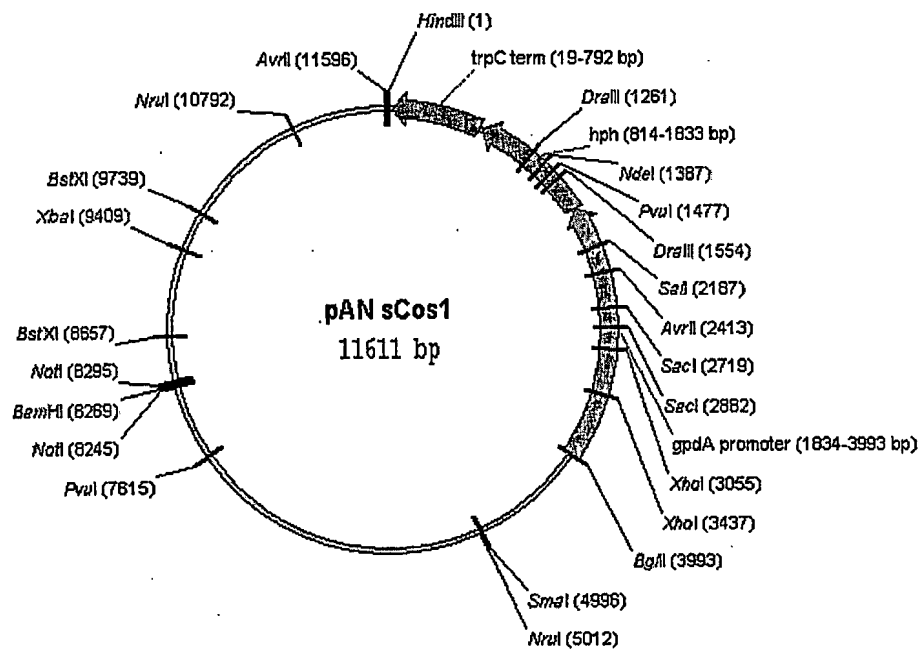
Fig. 1: Vektor pANsCos1

Fig. 2: Vektor pBinAHyg

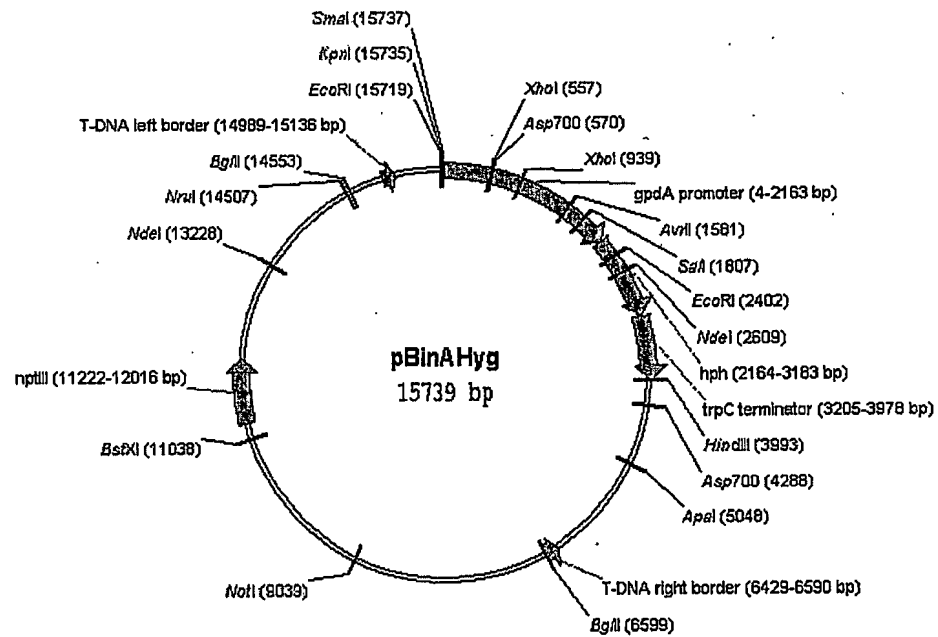


Fig. 3: Gels des Ergebnis einer PCR

Spur:

1 2 3 4 5 6 7 8 9



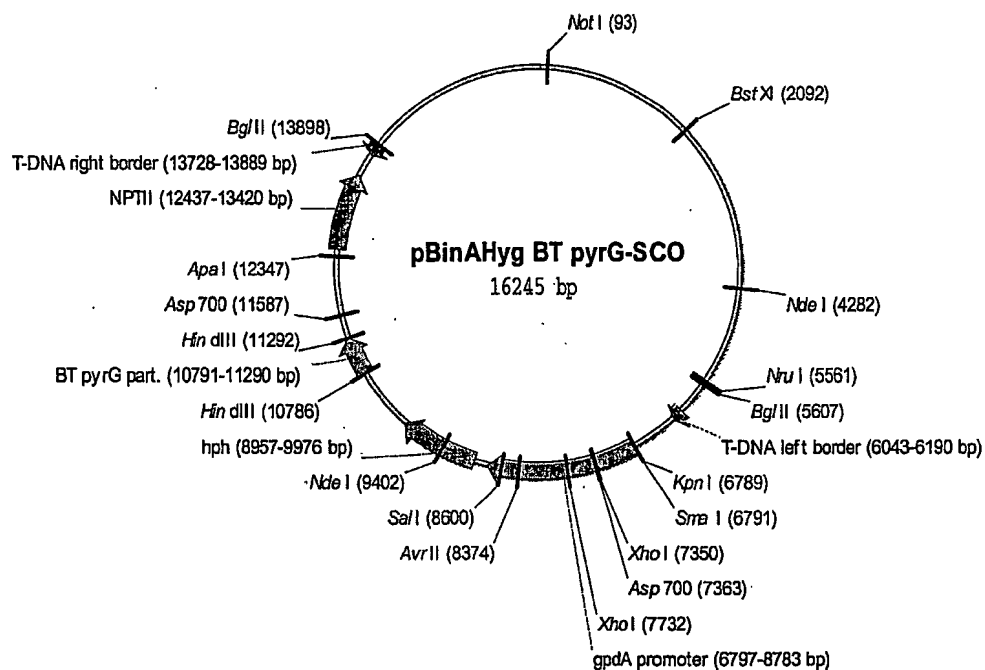
Fig. 4: Plasmid pBinAHyg BT pyrG-SCO

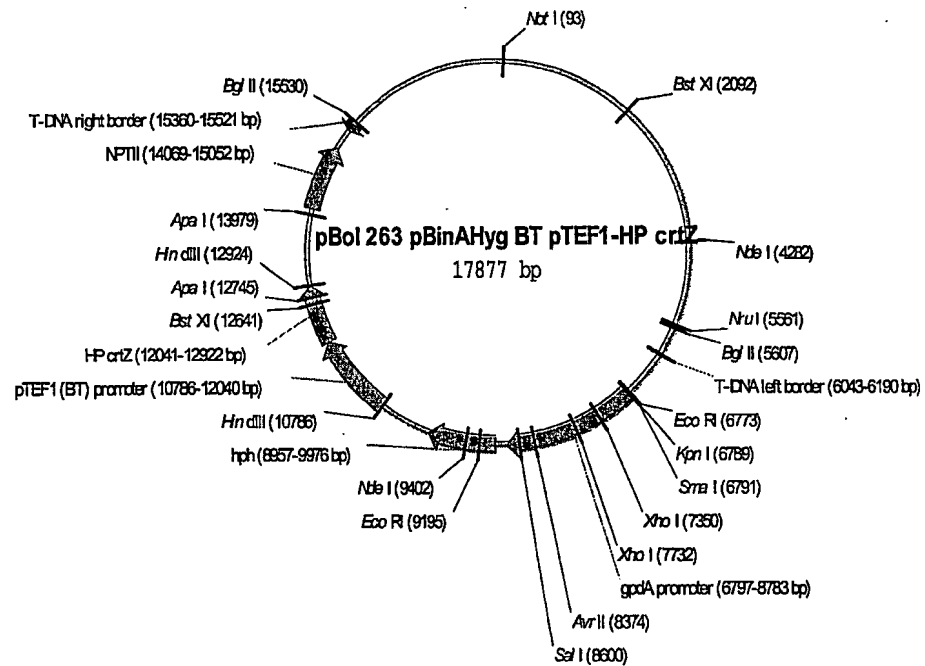
Fig. 5: Plasmid pBinAHygBTpTEF1-HPcrtZ

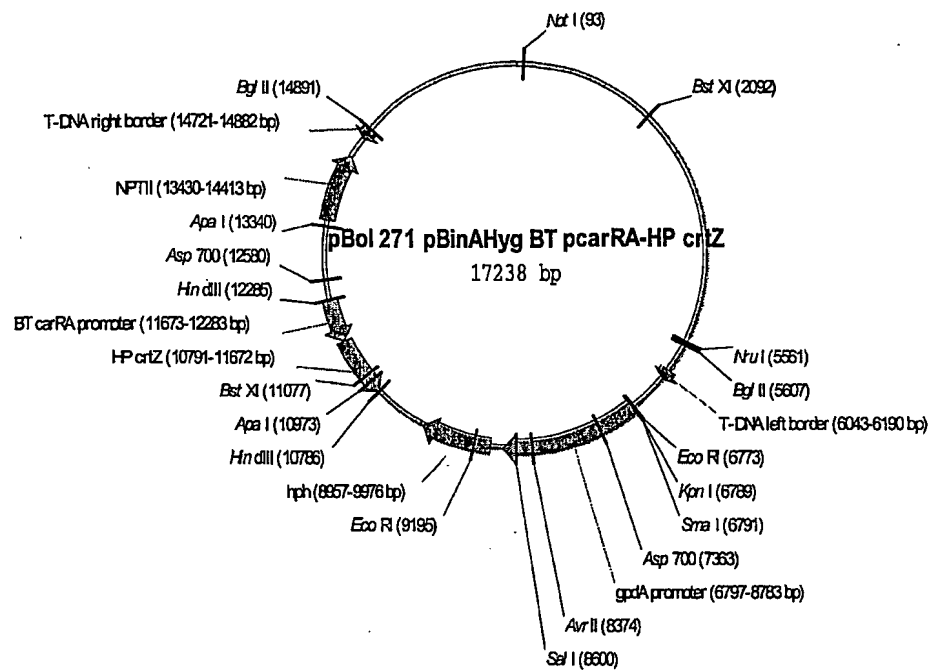
Fig. 6: Plasmid pBinAHyg-BTpcarRA-HPcrtZ

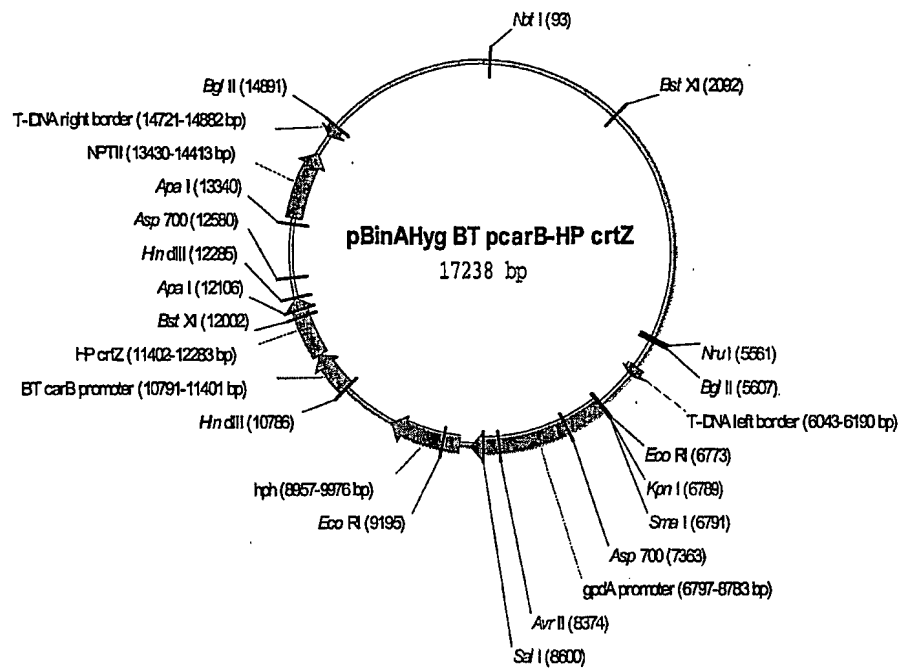
Fig. 7: Plasmid pBinAHygBTpcarB-HPcrtZ

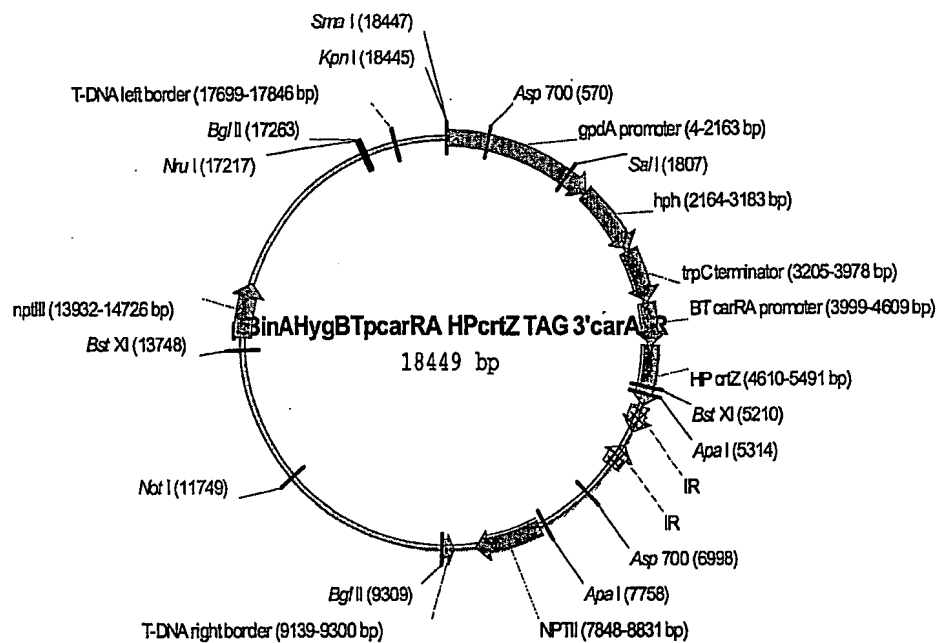
Fig. 8: Plasmid p-carRA-HPcrtZ-TAG-3'carA-IR

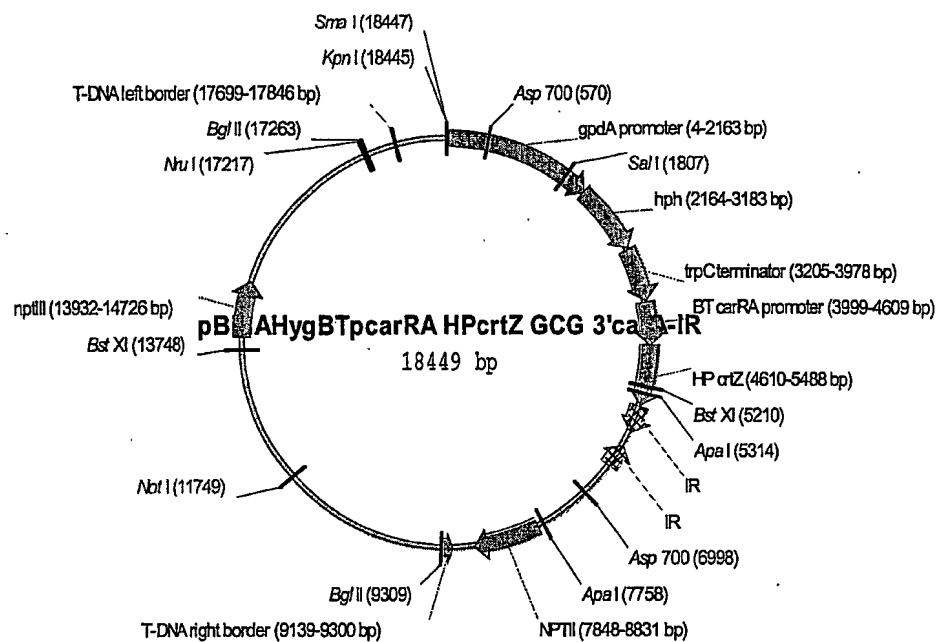
Fig. 9: Plasmid p-carRA-HPcrtZ-GCG-3'carA-IR

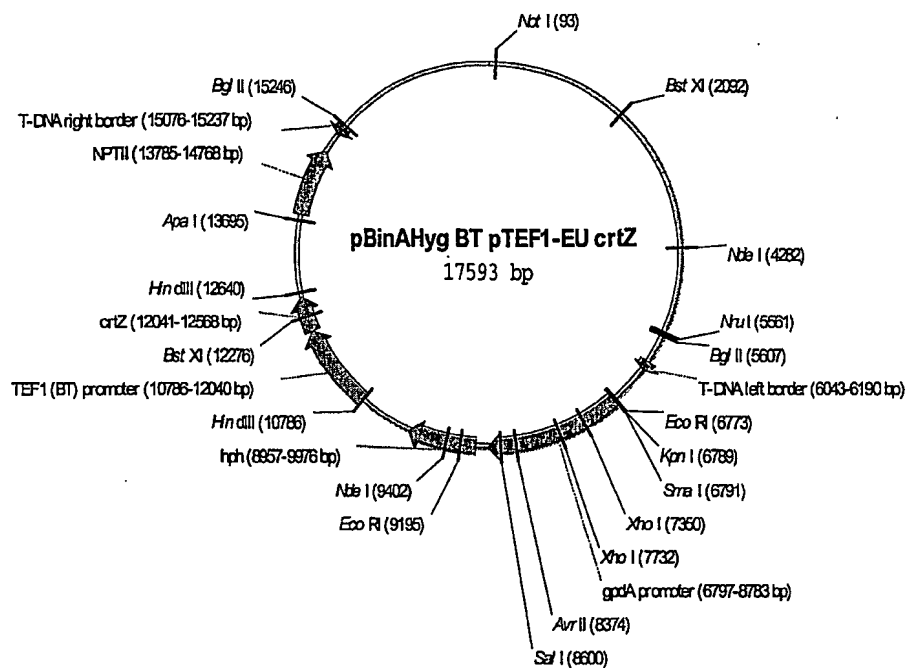
Fig. 10: Plasmid pBinAHygBTpTEF1-EUcrtZ

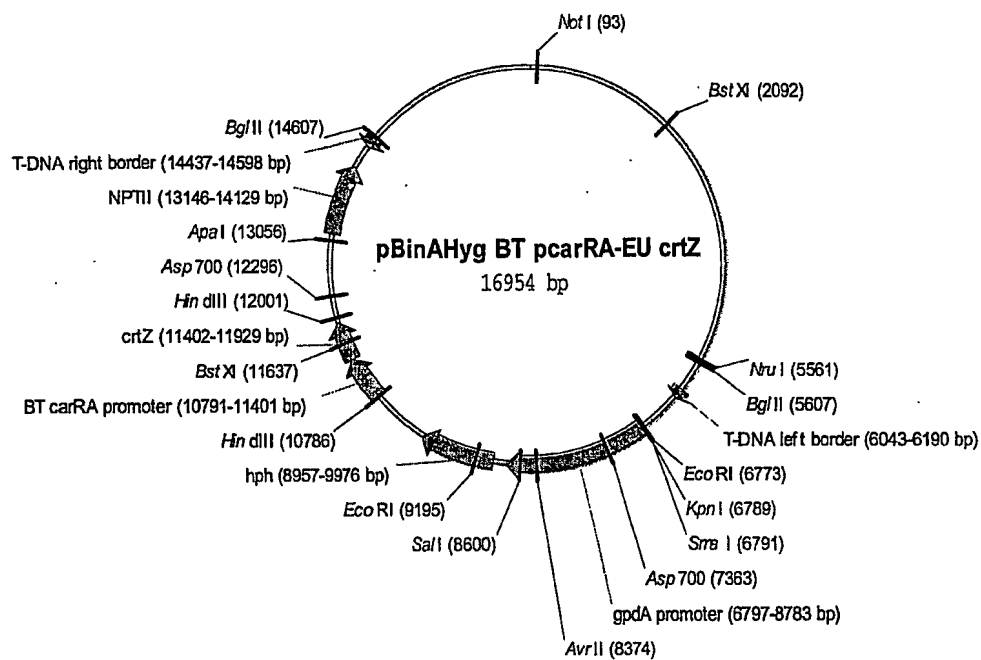
Fig. 11: Plasmid pBinAHygBTpcarRA-EUcrtZ

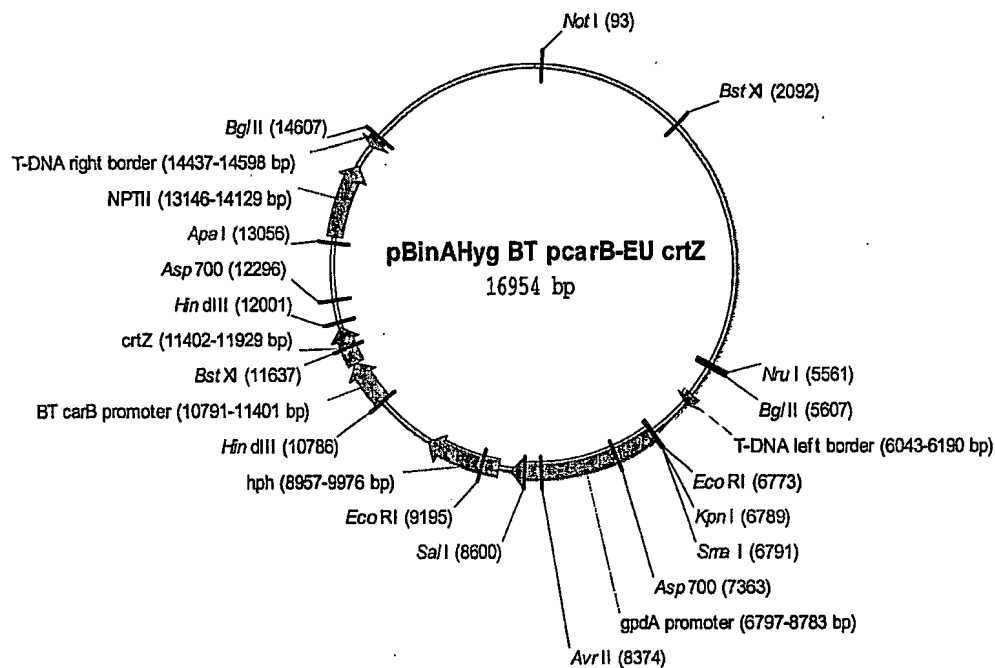
Fig. 12: Plasmid pBinAHygBTpcarB-EUcrtZ

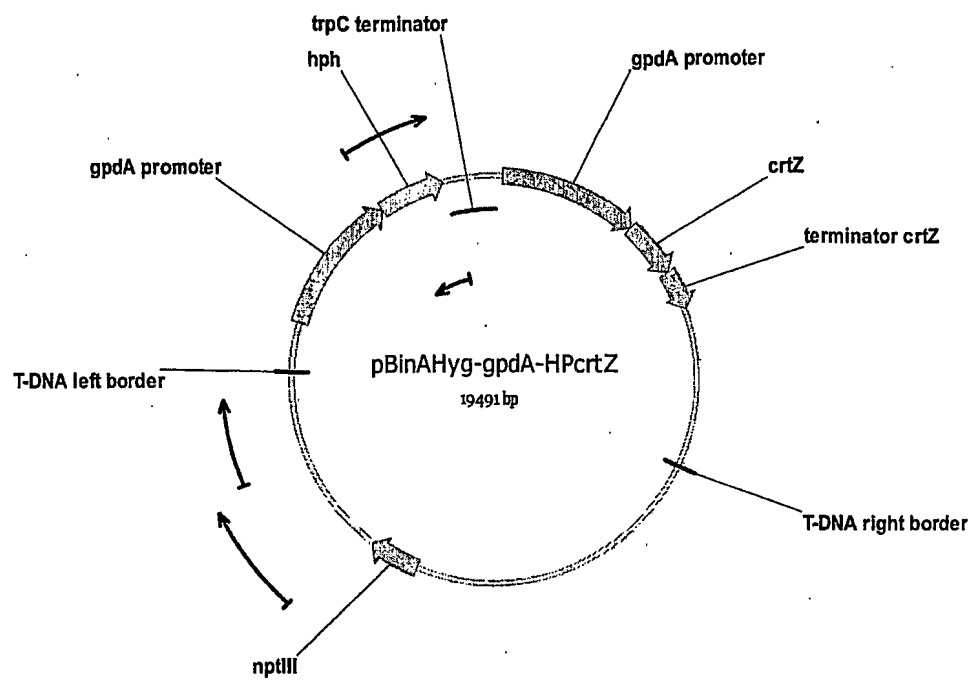
Fig. 13: Plasmid p-BinAHyg-gpdA-HPcrtZ

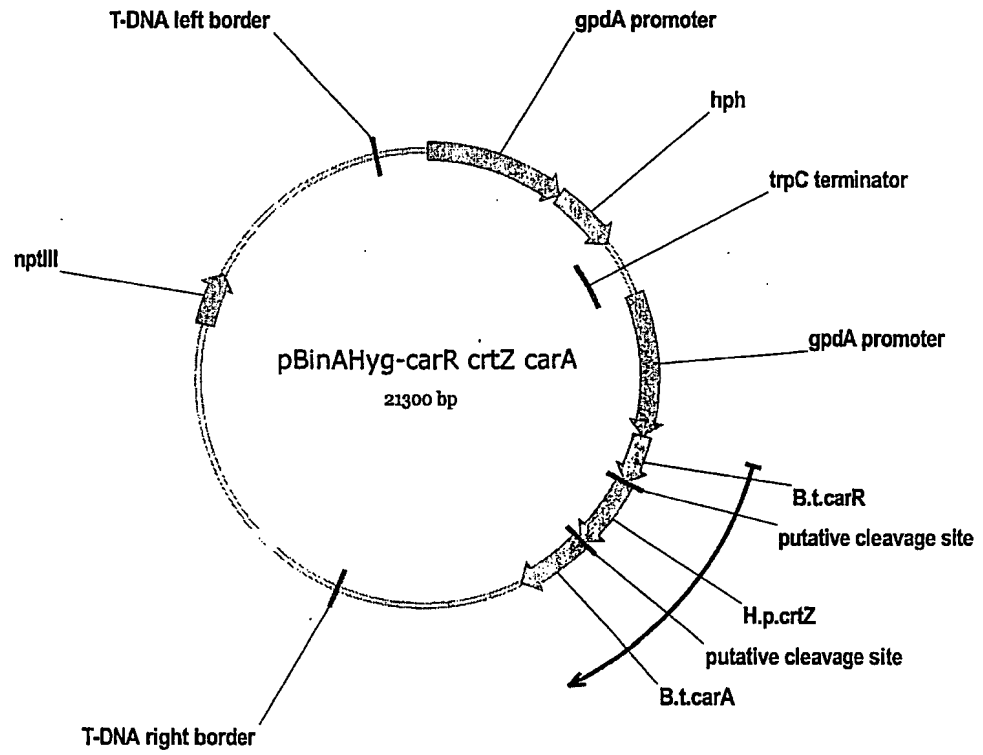
Fig. 14: Plasmid pBinAHyg-carRcrtZcarA

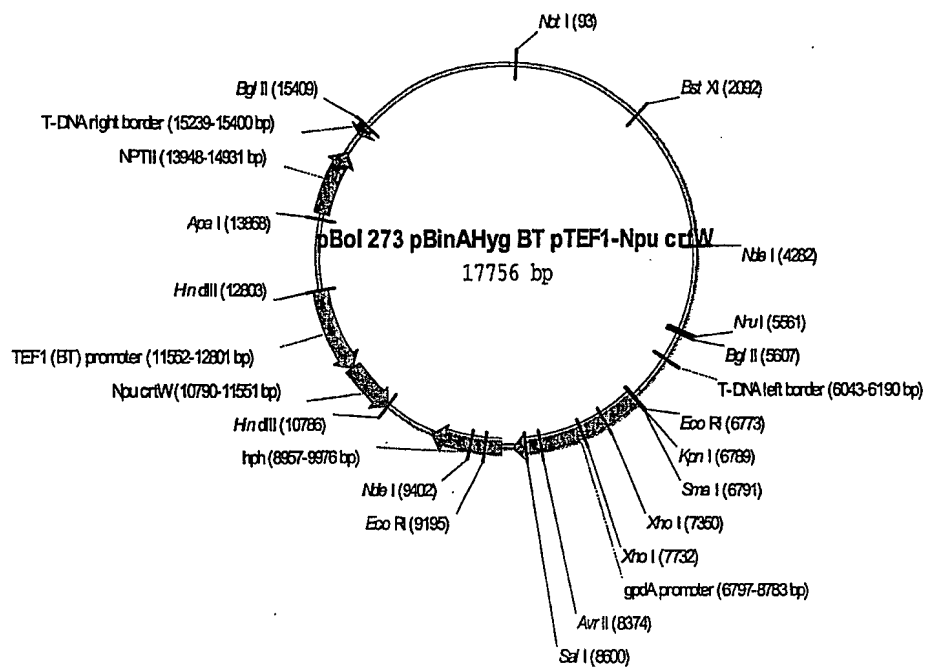
Fig. 15: Plasmid pBinAHyg-BTpTEF1-NPcrtW

Fig. 16: Plasmid pBinAHyg_BTpcarRA_NPcrtW

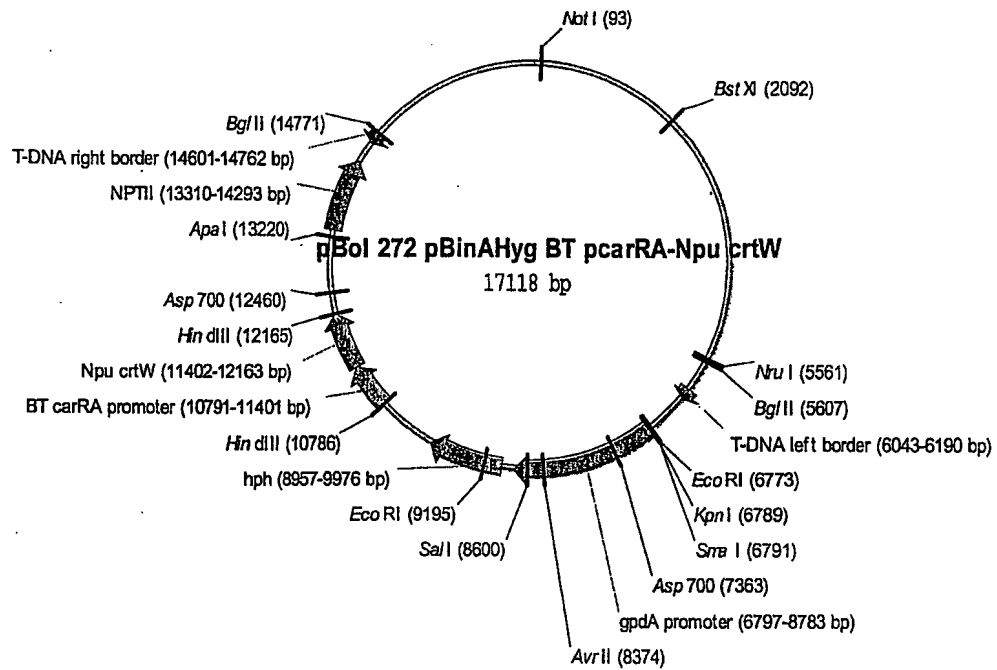


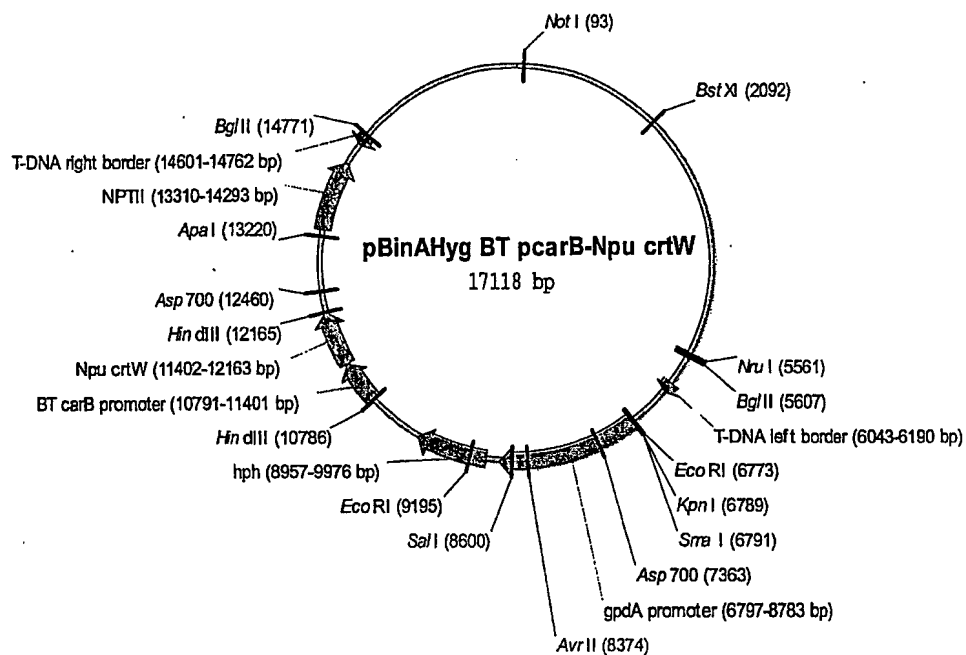
Fig. 17: Plasmid pBinAHyg-BTpcarB-NPcrtW

Fig. 18: Plasmid pBinAHygBTpcarRA-HPcrtZ-BTpcarRA-NpucrtW

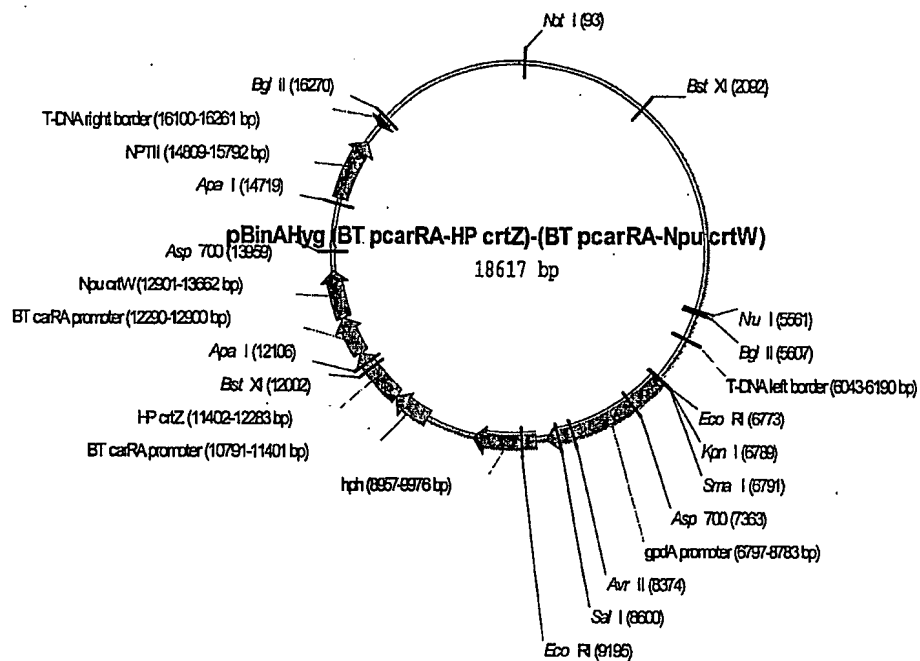


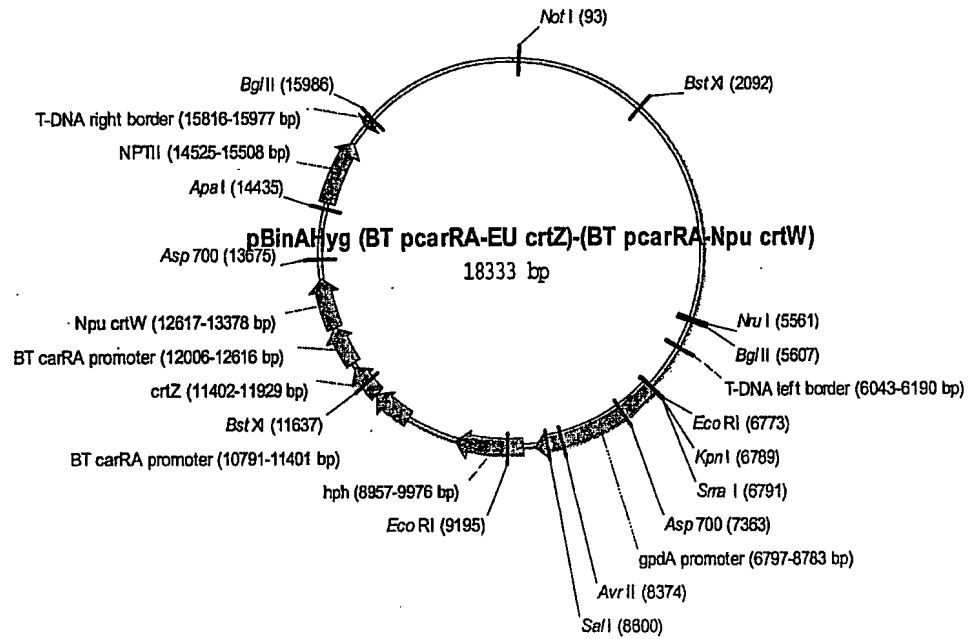
Fig. 19: Plasmid pBinAHygBTpcarRA-EUcrtZ-BTpcarRA-NpucrtW

Fig. 20: carB

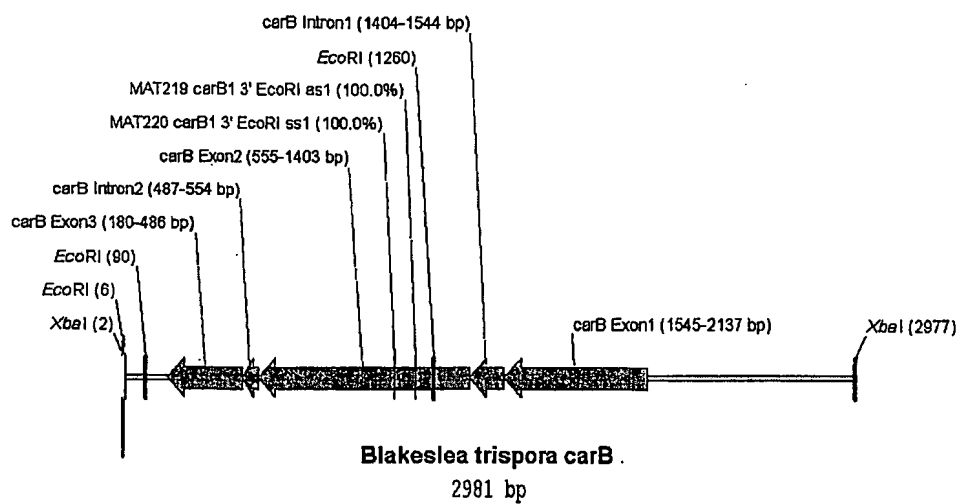


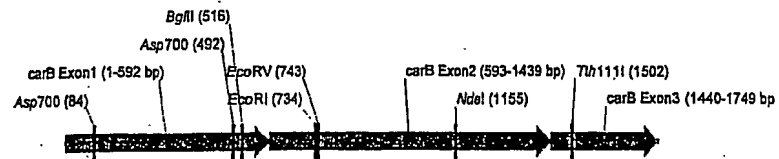
Fig. 21: CDS von carB

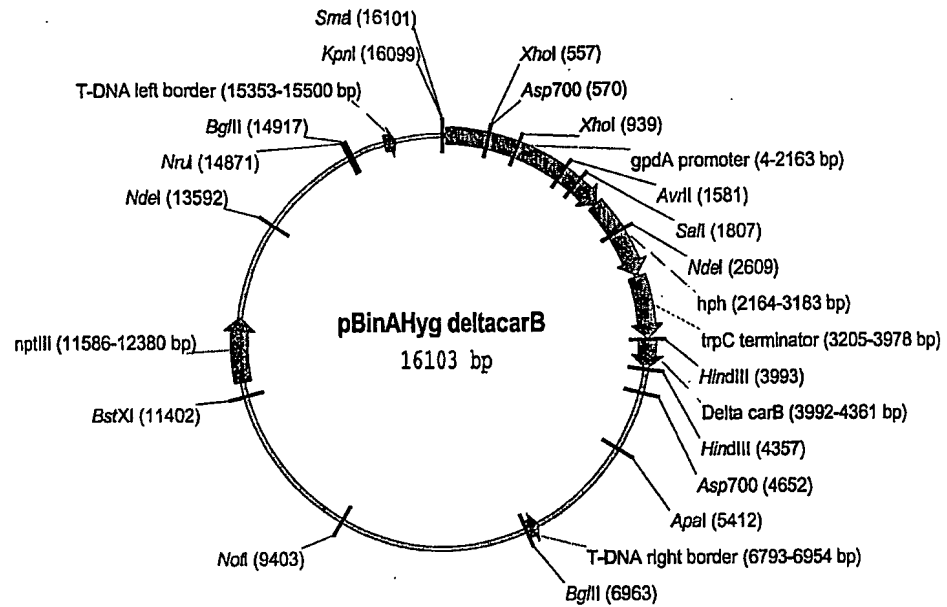
Fig. 22: Vektor pBinAHyg Δ carB

Fig. 23: HPLC Standard

Sequence File : C:\HPCHEM\2\SEQUENCE\CAROTIN.S
Method : C:\HPCHEM\2\METHODS\CAROTIN.M
Last changed : 11/12/2003
Carotinoide Nucleosil 100-7, C18 250*4, 0mm

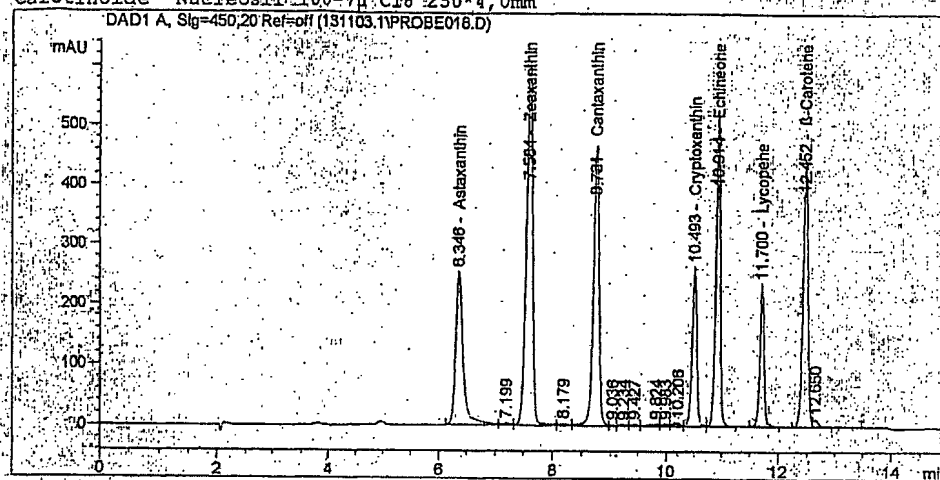
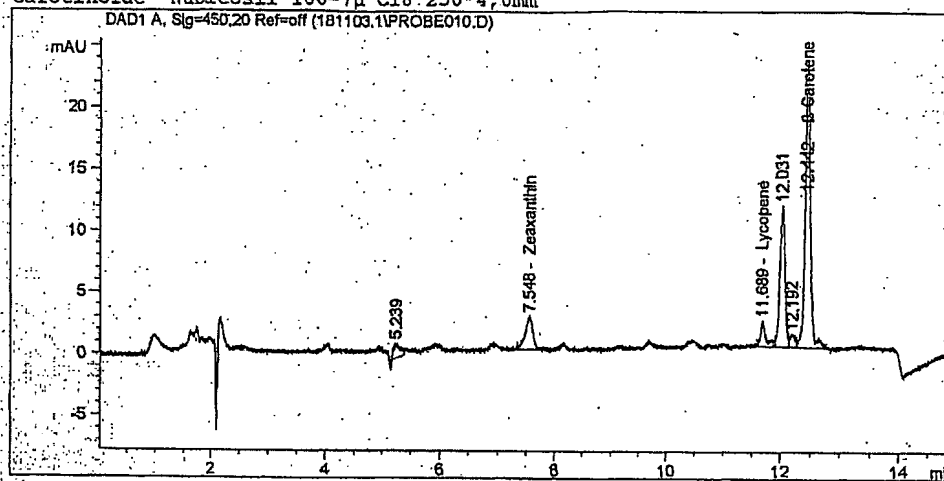


Fig. 24: HPLC

Sequence File : C:\HPCHEM\2\SEQUENCE\CAROTIN.S
Method : C:\HPCHEM\2\METHODS\CAROTIN.M
Last changed : 11/18/2003
Carotinoide Nucleosil 100-7µ C18 250*4, 0mm



SEQUENCE LISTING

<110> BASF AG

<120> Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen
mittels gentechnisch veränderter Organismen der Gattung
Blakeslea, mit dem Verfahren hergestellte Carotinoide oder deren
Vorstufen und deren Verwendung

<130> BASF/NAE877/03

<160> 80

<170> PatentIn version 3.2

<210> 1

<211> 2160

<212> DNA

<213> Artificial

<220>

<223> Promotor

<400> 1

ctttcgacac tgaaatacgt cgagcctgct ccgcttggaa gggcgagga gcctcgtcct	60
gtcacaacta ccaacatgga gtacgataag ggccagttcc gccagctcat taagagccag	120
ttcatgggcg ttggcatgat ggccgtcatg catctgtact tcaagtacac caacgctcct	180
ctgatccagt cgatcatccg ctgaaggcgc tttcgaatct ggttaagatc cacgtcttcg	240
ggaagccagc gactggtgac ctccagcgtc cctttaaggc tgccaacagc tttctcagcc	300
agggccagcc caagaccgac aaggcctccc tccagaacgc cgagaagaac tggaggggtg	360
gtgtcaagga ggagtaagct ccttattgaa gtcggaggac ggagcgggtg caagaggata	420
ttcttcgact ctgtattata gataagatga tgaggaattg gaggtagcat agcttcattt	480
ggatttgctt tccaggctga gactctagct tggagcatag agggtccttt ggctttcaat	540

atttctcaagt atctcgagtt tgaacttatt ccctgtgaac cttttattca ccaatgagca . 600

ttggaatgaa catgaatctg aggactgcaa tcgccatgag gttttcgaaa tacatccgga . 660

tgtcgaaggc ttggggcacc tgcgttggtt gaatttagaa cgtggcacta ttgatcatcc 720

gatagctctg caaagggcgt tgcacaatgc aagtcaaacg ttgctagcag ttccaggtgg 780

aatgttatga tgagcattgt attaaatcag gagatatagc atgatctcta gttagctcac 840

cacaaaagtc agacggcgta accaaaagtc acacaacaca agctgtaagg atttcggcac 900

ggctacggaa gacggagaag ccaccttcag tggactcgag taccatttaa ttctatttgt 960

gtttgatcga gacctaatc agcccctaca acgaccatca aagtcgtata gctaccagtg 1020

aggaagtgga ctcaaatcga cttcagcaac atctcttgga taaactttaa gcctaaacta 1080

tacagaataa gatagggtgga gagcttatac cgagctccca aatctgtcca gatcatgggtt 1140

gaccggtgcc tggatcttcc tatagaatca tccttattcg ttgacctagc tgattctgga 1200

gtgaccaga gggatcatgac ttgagcctaa aatccgccgc ctccaccatt thtagaaaaa 1260

tgtgacgaac tcgtgagctc tgtacagtga ccggtgactc tttctggcat gcggagagac 1320

ggacggacgc agagagaagg gctgagtaat aagccactgg ccagacagct ctggcggctc 1380

tgaggtgcag tggatgatta ttaatccggg accggccgcc cctccgcccc gaagtggaaa 1440

ggctggtgtg cccctcgttg accaagaatc tattgcatca tcggagaata tggagcttca 1500

tcgaatcacc ggagtaagc gaaggagaat gtgaagccag ggggtgtatag ccgtcggcga 1560

aatagcatgc cattaaccta ggtacagaag tccaattgct tccgatctgg taaaagattc 1620

acgagatagt accttctccg aagtaggtag agcgagtacc cggcgcgtaa gctccctaat 1680

tggcccatcc ggcattctgta gggcgtccaa atatcgtgcc tctcttgctt tgcccgggtg 1740

atgaaaccgg aaaggccgct caggagctgg ccagcggcgc agaccgggaa cacaagctgg 1800

cagtcgaccc atccgggtgct ctgcactcga cctgctgagg tccctcagtc cctggtaggc 1860
agctttgccc cgtctgtccg cccgggtgtgt cggcgggggtt gacaagggtcg ttgcgtcagt 1920
ccaacatttg ttgcatatt ttctgtctct ccccaccagc tgctcttttc tttctctttt 1980
cttttcccat cttcagtata ttcatcttcc catccaagaa cctttatttc ccctaagtaa 2040
gtactttgct acatccatac tccatccttc ccatccctta ttcctttgaa cctttcagtt 2100
cgagctttcc cacttcacgc cagcttgact aacagctacc ccgcttgagc agacatcacc 2160

<210> 2

<211> 774

<212> DNA

<213> Artificial

<220>

<223> Terminator

<220>

<221> misc_feature

<222> (267)..(267)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (475)..(475)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (566)..(566)

<223> n is a, c, g, or t

<400> 2

cgatccactt aacgttactg aaatcatcaa acagcttgac gaatctggat ataagatcgt 60
tggtgtcgat gtcagctccg gagttgagac aaatgggtgtt caggatctcg ataagatacg 120

ttcatttgtc caagcagcaa agagtgcctt ctagtgattt aatagctcca tgtcaacaag 180
aataaaacgc gttttcgggt ttacctcttc cagatacagc tcatctgcaa tgcattaatg 240
cattgactgc aacctagtaa cgccttnacg gctccggcga agagaagaat agcttagcag 300
agctattttc attttcggga gacgagatca agcagatcaa cggtcgtcaa gagacctacg 360
agactgagga atccgctctt ggctccacgc gactatatat ttgtctctaa ttgtactttg 420
acatgctcct cttctttact ctgatagctt gactatgaaa attccgtcac cagcncctgg 480
gttcgcaaag ataattgcat gtttcttctt tgaactctca agcctacagg acacacattc 540
atcgtaggta taaacctega aatcanttcc tactaagatg gtatacaata gtaaccatgc 600
atggttgcct agtgaatgct ccgtaacacc caatacgccg gccgaaactt ttttacaact 660
ctcctatgag tcgtttaccc agaatgcaca ggtacacttg tttagaggta atccttcttt 720
ctagctagaa gtctctgtgt actgtgtaag cgcccactcc acatctccac tcga 774

<210> 3

<211> 15739

<212> DNA

<213> Artificial

<220>

<223> Vector

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 3

gatctttcga cactgaaata cgtcgagcct gtcgcgcttg gaagcggcga ggagcctcgt	60
cctgtcacia ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc	120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct	180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct	240
tcggaagcc agcgactggg gacctccagc gtccctttta ggctgccaac agctttctca	300
gccagggcca gccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg	360
gtggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttggaattg ctttccaggc tgagactcta gcttgagca tagaggggcc tttggctttc	540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga	600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgctt gttgaattta gaacgtggca ctattgatca	720
tccgatagct ctgcaaaggc cgttgcaaaa tgcaagtcaa acgttgctag cagttccagg	780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840
caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg	900
cacggctacg gaagacggag aagccacctt cagtggactc gaggaccatt taattctatt	960
tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca	1020
gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa	1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140
gttgaccggt gcctggatct tccatagaa tcatccttat tcgttgacct agctgattct 1200
ggagtgacct agagggtcat gacttgagcc taaaatccgc cgcctccacc attttagaa 1260
aaatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga 1320
gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380
ctctgagggtg cagtggatga ttattaatcc gggaccggcc gccctccgc ccgaagtgg 1440
aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcatcggaga atatggagct 1500
tcatcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560
cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgac tcgtaaaaga 1620
ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680
aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccg 1740
tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800
tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860
ggcagctttg ccccgctctgt ccgcccgggtg tgcggcgggg gttgacaagg tcgttgcgtc 1920
agtccaacat ttgttgccat attttctgc tctcccacc agctgctctt ttcttttctc 1980
tttcttttcc catcttcagt atattcatct tcccatcaa gaacctttat ttcccctaag 2040
taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100
gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160
accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
gtctccgacc tgatgcagct ctccgagggc gaagaatctc gtgctttcag cttcgatgta 2280
ggagggcggtg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgct tgacattggg 2400

gaattcagcg agagcctgac ctattgcac tcccgccgtg cacaggggtg cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctggt ctgcagccgg tcgcggaggc catggatgcg 2520

atcgctgcgg ccatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580

ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640

tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700

atgctttggg ccgaggactg ccccgaaagtc cggcacctcg tgcacgcgga ttccggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820

ttcggggatt cccaatacga ggtcgccaac atctttctct ggaggccgtg gttggcttgt 2880

atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgcagg atcgccgcgg 2940

ctccgggcgt atatgctccg cattggctct gaccaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggc cgatgcgacg caatcgccg atccggagcc 3060

gggactgtcg ggcgtacaca aatcgccgc agaagcgcgg ccgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cgcccagca ctcgccgag ggcaaaggaa 3180

tagagtagat gccgaccg cgatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300

tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagt ccttctagt 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaacct gtaacgcct ncaggctccg 3480

gcgaagagaa gaatagctta gcagagctat ttctattttc gggagacgag atcaagcaga 3540

tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggctcc acgcgactat 3600

atatttgtct ctaattgtac ttgacatgc tcctcttctt tactctgata gcttgactat 3660
gaaaattcog tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720
ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttcctactaa 3780
gatggtatac aatagtaacc atgcatgggt gcctagttaa tgcctcgtaa cacccaatac 3840
gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacagggtaca 3900
cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960
ctccacatct ccactcgacc tgcaggcatg caagcttggc gtaatcatgg tcatagctgt 4020
ttcctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc ggaagcataa 4080
agtgtaaagc ctgggggtgc taatgagtga gctaactcac attaattgcg ttgcgctcac 4140
tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 4200
cggggagagg cggtttgcgt attgggcaa agacaaaagg gcgacattca accgattgag 4260
ggagggaagg taaatattga cggaaattat tcattaaagg tgaattatca ccgtcaccga 4320
cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa 4380
ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta gcgacagaat 4440
caagtttgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg gtcatagccc 4500
ccttatttagc gtttgccatc ttttcatat caaaatcacc ggaaccagag ccaccaccgg 4560
aaccgcctcc ctgagagcgc ccaccctcag aaccgccacc ctgagagcca ccaccctcag 4620
agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggc ccgatctagt 4680
aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat tttgttttct 4740
atcgcgatatt aaatgtataa ttgcgggact ctaatcataa aaaccatct cataaataac 4800
gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata 4860

atcatcgcaa gaccggcaac aggattcaat ctttaagaaac tttattgcca aatgtttgaa 4920
cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga acgcagcaag 4980
atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt gatgtggacg 5040
ccgggcccga tcatattgtc gctcaggatc gtggcggtgt gcttgtcggc cgttgctgtc 5100
gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc gaagaactcc 5160
agcatgagat ccccgcgctg gaggatcatc cagccggcgt cccggaaaac gattccgaag 5220
cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag gttgggcgtc 5280
gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca agaaggcgat 5340
agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag 5400
cccattcgcc gccaaactct tcagcaatat cacgggtagc caacgctatg tcttgatagc 5460
ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca ttttccacca 5520
tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg tcgggcatgc 5580
gcgcttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct tcgtccagat 5640
catcctgata gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg 5700
cttggtggtc gaatgggcag gtagccggat caagcgatg cagccgccgc attgcatcag 5760
ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc tgccccggca 5820
cttcgcccga tagcagccag tcccttcccg cttcagtgc aacgtcgagc acagctgcgc 5880
aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc agttcattca 5940
gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga 6000
acacggcggc atcagagcag ccgattgtct gttgtgccca gtcatagccg aatagcctct 6060
ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgatc 6120

cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag 6180

gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg 6240

accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt agctcattaa 6300

actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca gttccaaacg 6360

taaaacggct tgtcccgct catcggcggg ggtcataacg tgactccctt aattctccgc 6420

tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg acaggatata 6480

ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcggatat ttaaaagggc 6540

gtgaaaaggt ttatccgttc gtccatttgt atgtgcatgc caaccacagg gttccccaga 6600

tctggcgccg gccagcgaga cgagcaagat tggccgcgc cggaaacgat ccgacagcgc 6660

gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca gaatgccata 6720

gtgggcggtg acgtcgttcg agtgaaccag atcgcgcagg aggcccgga gcaccggcat 6780

aatcaggccg atgccgacag cgtcgagcgc gacagtgttc agaattacga tcaggggtat 6840

gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa cgcgcggatt 6900

ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg 6960

acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa cgaggtcggc 7020

gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca gccggcgctt 7080

tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc catgctggcg 7140

gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt tctgatcggg 7200

aatgcccga gcttcaggca ggcgtgtc gcctaccgag atggcgcgcg catccatgcc 7260

ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg cttectctgc 7320

gaggcgggtt tttcggccgg ggacgcgctc aatgcgctga tgacaatcag ctacttcact 7380

gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg cggcggcacc 7440
gttgaacagg ctccgctctc gccgctgttg cgggccgcga tagacgcctt cgacgaagcc 7500
ggtcgggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg 7560
aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc aggaccgctg 7620
ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct ccccctttcc 7680
accgcgtcag acgcccgtag cagcccgtta cgggcttttt catgccctgc cctagcgtcc 7740
aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc gcttcctcgc 7800
tactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 7860
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 7920
gccagcaaaa gcccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggetcc 7980
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcacag 8040
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 8100
ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgcttttcc 8160
getgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc atcctttttc 8220
gcacgatata caggattttg ccaaaggggt cgtgtagact ttccttggtg tatccaacgg 8280
cgtcagccgg gcaggatagg tgaagtaggc ccaccgcga gcgggtgttc cttcttctact 8340
gtcccttatt cgcacctggc ggtgctcaac ggaatcctg ctctgcgagg ctggccggct 8400
accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag 8460
ggcagcccac ctatcaagggt gtactgcctt ccagacgaac gaagagcgat tgaggaaaag 8520
gcggcgggcg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca gggctacaaa 8580
atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa tggcgacctg 8640

ggcgcctgg gggcctgct gaaactctgg ctcaccgacg acccgcgac ggcgcggttc 8700

ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga cgagcttggc 8760

aaggtcatga tgggcgtggt ccgcccagg gacagagccat gactttttta gccgctaaaa 8820

cggcggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca agaagagcga 8880

cttcgcggag ctggtgaagt acatcaccca cgagcaaggc aagaccgagc gcctttgcga 8940

cgctcaccgg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc cctgcaaacg 9000

cgcagaaac gccgtcgaag ccgtgtgcga gacaccgcg ccgcccggcgt tgtggatacc 9060

tcgcggaaaa cttggccctc actgacagat gagggggcga cgttgacact tgaggggccg 9120

actcaccgg cgcgcgcttg acagatgagg ggcaggctcg atttcggccg gcgacgtgga 9180

gctggccagc ctgcgaaatc ggcgaaaacg cctgatttta cgcgagtttc ccacagatga 9240

tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc gcgactactg 9300

acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga tgaggggcgc 9360

acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc aagggtttcc 9420

gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca atatttataa 9480

accttgtttt taaccagggc tgccctctgt gcgcgtgacc gcgcacgccg aaggggggtg 9540

cccccccttc tcgaaccctc ccggcccgt aacgcgggac tcccatcccc ccaggggctg 9600

cgccctcgg ccgcgaacgg cctcaccca aaaatggcag cgctggcagt ccttgccatt 9660

gccgggatcg gggcagtaac gggatgggcg atcagcccga gcgcgacgcc cggaagcatt 9720

gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag tgagggcggc 9780

ggcctgggtg gggcctgcc cttcacttcg gccgtgggg cattcacgga cttcatggcg 9840

gggcgggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc cgtgctcgtg 9900

ttcgggggtg cgataaaccc agcgaaccat ttgaggtgat aggtaagatt ataccgaggt 9960
atgaaaacga gaattggacc ttacagaat tactctatga agcgccatat ttaaaaagct 10020
accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat attgacaata 10080
ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga tttcaggggg 10140
caaggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca taaaaacttg 10200
catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt ctatcataat 10260
tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgccg gatgactttg 10320
tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg tgccagggtgc 10380
tgccctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct gattacgtgc 10440
agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca taccaccagc 10500
tcaaaggggtg acagcaggct cataagacgc ccagcgtcg ccatagtgcg ttcaccgaat 10560
acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca gcgctggcgc 10620
gatttagccc cgacatagcc ccactgttcg tccatttcgg cgagacgat gacgtcactg 10680
cccggctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga cgtaaaatcg 10740
tgttgaggcc aacgccata atgcgggctg ttgccggca tccaacgcca ttcattggca 10800
tatcaatgat tttctggtgc gtaccgggtt gagaagcggg gtaagtgaac tgcagttgcc 10860
atgttttacg gcagtgcgag cagagatagc gctgatgtcc ggcgggtgctt ttgccgttac 10920
gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa gccactggag 10980
cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc cataattgtg 11040
gtttcaaaaat cggctccgct gatactatgt tatacgccaa ctttgaaaac aactttgaaa 11100
aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg gagttcgtct 11160

tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaaataata 11220

aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt 11280

aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag aaaatgaaaa 11340

cctatatatta aaaatgacgg acagccggta taaagggacc acctatgatg tggaacggga 11400

aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtcc tgcactttga 11460

acggcatgat ggctggagca atctgctcat gagtaggcc gatggcgtcc tttgctcgga 11520

agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg agtgcacag 11580

gctctttcac tccatcgaca tatcggttg tccctatacg aatagcttag acagccgctt 11640

agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg aaaactggga 11700

agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaaaagcc 11760

cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct ttgtgaaaga 11820

tggtcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca agtggtatga 11880

cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt atgtcgagct 11940

atTTTTtgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt atattttact 12000

ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca 12060

ccgacttctt ccgcatcaag tgTTTTggct ctcaggccga ggcccacggc aagtatttgg 12120

gcaaggggtc gctggtattc gtgcaggga agattcgaa taccaagtac gagaaggacg 12180

gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg gacaccaagg 12240

caccagggcg gtcaaatcag gaataagggc acattgcccc ggcgtgagtc ggggcaatcc 12300

cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa gaactgatcg 12360

acgcggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc atgcgtgcgc 12420

cccgcgaaac cttccagtcg gtcggctcga tgggtccagca agctacggcc aagatcgagc 12480
gcgacagcgt gcaactggct cccctgccc tgcccgcgc atcggccgcc gtggagcgtt 12540
cgcgctgtct cgaacaggag gggcagggtt tggcgaagtc gatgaccatc gacacgcgag 12600
gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa caggtcagcg 12660
aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa atgcagcttt 12720
ccttgttcga tattgcgccg tggccggaca cgatgcgagc gatgccaac gacacggccc 12780
gctctgcctt gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg caaaacaagg 12840
tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgctcgag ctgcggggccg 12900
acgatgacga actggtgtgg cagcagggtg tggagtacgc gaagcgcacc cctatcggcg 12960
agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggtcg atcaatggcc 13020
ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg atgggcttca 13080
cgtccgaccg cgttgggcac ctggaatcgg tgtoctgct gcaccgcttc cgcgtcctgg 13140
accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc gtcgtgctgt 13200
ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg tcgccgacgg 13260
cccgcaggat gttcgactat ttcagctcgc accgggagcc gtaccgcctc aagctggaaa 13320
ccttcgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc gagcaggctg 13380
gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg gtcaatgatg 13440
acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg ggttcagcag 13500
ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact tgcttcgctc 13560
agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag gattaaaatt 13620
gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc aggatttccg 13680

cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg tttacgagca 13740
cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg tggcattcgg 13800
cgcctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg acggcccaa 13860
ggacgctcac aaggcgcac tgtccggcgt tttcgtggag cccgaacagc gaggccgagg 13920
ggtcgccggt atgctgctgc gggcgttgcc ggcgggttta ttgctcgtga tgatcgtccg 13980
acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac ttaatatctc 14040
gctattctgg agcttggtgt ttatttcggt ctaccgcctg ccgggcgggg tcgcgcgac 14100
ggtaggcgct gtgcagccgc tgatggctgt gttcatctct gccgtctgc taggtagccc 14160
gatacgattg atggcgggcc tgggggctat ttgcggaact gcgggcgtgg cgctgttggt 14220
gttgacacca aacgcagcgc tagatcctgt cggcgctcga gcgggcctgg cgggggcggt 14280
ttccatggcg ttcggaaccg tgctgaccgc caagtggcaa cctcccgctc ctctgctcac 14340
ctttaccgpc tggcaactgg cggcgggagg acttctgctc gttccagtag ctttagtggt 14400
tgatccgcc aatccgatgc ctacaggaac caatgttctc ggcttggcgt ggctcggcct 14460
gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac tcgaacctac 14520
agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc cggggatgca 14580
tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtgag caatggatag 14640
gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc ttcctcagcg 14700
gctttatcca gcgatttcct attatgtcgg catagtcttc aagatcgaca gcctgtcacg 14760
gttaagcgag aatgaataa gaaggctgat aatcggatc tctgcgaggg agatgatatt 14820
tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct ccgcgagatc 14880
atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc ggtaacatga 14940

gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact gatgggctgc 15000
ctgtatcgag tggatgattt gtgccgagct gccggtcggg gagctggttg ctggctgggtg 15060
gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac acattgcgga 15120
cgtttttaat gtactggggg ggtttttctt ttcaccagt agacgggcaa cagctgattg 15180
cccttcaccg cctggccctg agagagttgc agcaagcggc ccacgctggt ttgccccagc 15240
aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat aaatcaaaag 15300
aatagcccga gatagggttg agtggtgttc cagtttgga caagagtcca ctattaaaga 15360
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc cactacgtg 15420
aaccatcacc caaatcaagt tttttggggc cgaggtgccg taaagcacta aatcggaacc 15480
ctaaagggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg 15540
aaggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg ggaagggcga 15600
tcgggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 15660
ttaagttggg taacgccagg gttttccag tcacgacgtt gtaaaacgac ggccagtga 15720
ttcgagctcg gtacccggg 15739

<210> 4

<211> 11611

<212> DNA

<213> Artificial

<220>

<223> Vector

<220>

<221> misc_feature

<222> (227)..(227)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (318)..(318)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (526)..(526)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (8946)..(8946)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10028)..(10028)

<223> n is a, c, g, or t

<400> 4

agcttgcattg cctgcagggtc gagtggagat gtggagtggg cgcttacaca gtacacgagg 60

acttctagct agaaagaagg attacctcta aacaagtgtg cctgtgcatt ctgggtaaac 120

gactcatagg agagttgtaa aaaagtttcg gccggcgat tggtgtgtac ggagcattca 180

ctaggcaacc atgcatgggtt actattgtat accatcttag taggaantga ttctgaggtt 240

tatacctacg atgaatgtgt gtctgttagg cttgagagtt caaggaagaa acatgcaatt 300

atctttgcga acccaggngc tgggtgacgga attttcatag tcaagctatc agagtaaaga 360

agaggagcat gtcaaagtac aattagagac aaatatatag tcgcgtggag ccaagagcgg 420

attcctcagt ctcttaggtc tcttgacgac cgttgatctg cttgatctcg tctcccga 480

atgaaaatag ctctgctaag ctattcttct ctctgccgga gcctgnaagg cgttactagg 540

ttgcagtcaa tgcattaatg cattgcagat gagctgtatc tggaagaggt aaacccgaaa 600

acgcgttttta ttcttggtga catggagcta ttaaatact agaaggcact ctttgctgct 660
tggacaaatg aacgtatctt atcgagatcc tgaacacccat ttgtctcaac tccggagctg 720
acatcgacac caacgatctt atatccagat tcgtcaagct gtttgatgat ttcagtaacg 780
ttaagtggat cgatcccgcg gtcggcatct actctattcc tttgcctcg gacgagtgct 840
ggggcgctcg tttccactat cggcgagtag ttctacacag ccatcggtcc agacggccgc 900
gcttctgcgg gcgatttggtg taagcccgac agtcccggt cccgatcgga cgattgcgtc 960
gcatcgaccc tgcgccaag ctgcatcacc gaaattgccg tcaaccaagc tctgatagag 1020
ttggtcaaga ccaatgcgga gcatatacgc ccggagccgc ggcgatcctg caagctccgg 1080
atgcctccgc tcgaagtagc ggtctgctg ctccatacaa gccaaaccag gcctccagaa 1140
gaagatgttg gcgacctgt attgggaatc cccgaacatc gcctcgctcc agtcaatgac 1200
cgctgttatg cggccattgt ccgtcaggac attgttgag ccgaaatccg cgtgcacgag 1260
gtgccggact tcggggcagt cctcgcccca aagcatcagc tcatcgagag cctgcgcgac 1320
ggacgcactg acggtgtcgt ccatcacagt ttgccagtga tacacatggg gatcagcaat 1380
cgcgcatatg aaatcacgcc atgtagtgtg ttgaccgatt ccttgccggtc cgaatgggcc 1440
gaaccgcgtc gtctggctaa gatcgccgc agcgatcgca tccatggcct ccgcgaccgg 1500
ctgcagaaca gcgggcagtt cggtttcagg cagggtcttg aacgtgacac cctgtgcacg 1560
gcgggagatg caataggtca ggctctcgct gaattcccca atgtcaagca cttccggaat 1620
cgggagcgcg gccgatgcaa agtgccgata aacataacga tctttgtaga aaccatcggc 1680
gcagctatctt acccgagga catatccag ccctcctaca tcgaagctga aagcacgaga 1740
ttcttcgccc tccgagagct gcatcaggtc ggagacgctg tcgaactttt cgatcagaaa 1800
cttctcgaca gacgtcgcg tgagttcagg catggtgatg tctgctcaag cggggtagct 1860

gttagtcaag ctgcgatgaa gtgggaaagc tcgaactgaa aggttcaaag gaataaggga 1920

tgggaaggat ggagtatgga tgtagcaaag tacttactta ggggaaataa aggttcttgg 1980

atgggaagat gaatatactg aagatgggaa aagaaagaga aaagaaaaga gcagctggtg 2040

gggagagcag gaaaatatgg caacaaatgt tggactgacg caacgacctt gtcaaccccg 2100

ccgacacacc gggcggacag acggggcaaa gctgcctacc agggactgag ggacctcagc 2160

aggtcgagtg cagagcaccg gatgggtcga ctgccagctt gtgttcccg tctgcgccgc 2220

tggccagctc ctgagcggcc tttccggttt catacaccgg gcaaagcagg agaggcacga 2280

tatttggacg ccctacagat gccggatggg ccaattaggg agcttacgcg ccgggtactc 2340

gctctaccta cttcggagaa ggtactatct cgtgaatctt ttaccagatc ggaagcaatt 2400

ggacttctgt acctaggta atggcatgct atttcgccga cggctataca cccttggtt 2460

cacattctcc ttcgcttact gccggtgatt cgatgaagct ccatattctc cgatgatgca 2520

atagattctt ggtcaacgag gggcacacca gcctttccac ttcggggcgg aggggcggcc 2580

ggccccgat taataatcat ccaactgcacc tcagagccgc cagagctgtc tggccagtgg 2640

cttattactc agcccttctc totggtccg tccgtctctc cgcagccag aaagagtcac 2700

cggtcactgt acagagctca cgagttcgtc acatTTTTtT acaaatgggtg gaggcggcgg 2760

attttaggct caagtcatga ccctctgggt cactccagaa tcagctaggt caacgaataa 2820

ggatgattct ataggaagat ccaggcaccg gtcaaccatg atctggacag atttgggagc 2880

tcggtataag ctctccacct atcttattct gtatagttta ggcttaaagt ttatccagga 2940

gatgttgctg aagtcgattt gagtccactt cctcactggt agctatacga ctttgatggt 3000

cgtttaggg gctgtattag gtctcgatca aacacaaata gaattaaatg gtactcgagt 3060

ccactgaagg tggcttctcc gtcttccgta gccgtgccga aatccttaca gcttgtgttg 3120

tgtgactttt gggtacgccg tctgactttt gtggtgagct aactagagat catgctatat 3180
ctcctgattt aatacaatgc tcatcataac attccacctg gaactgctag caacgtttga 3240
cttgacattgt gcaacgccct ttgcagagct atcggatgat caatagtgcc acgttctaaa 3300
ttcaaccaac gcaggtgccc caagccttcg acatccggat gtatttcgaa aacctcatgg 3360
cgattgcagt cctcagattc atgttcattc caatgctcat tggngaataa aagggtcaca 3420
gggaataagt tcaaactcga gatacttgag aatattgaaa gccaaaggac cctctatgct 3480
ccaagctaga gtctcagcct ggaaagcaaa tccaaatgaa gctatgctac ctccaattcc 3540
tcatcatctt atctataata cagagtogaa gaatatcctc ttgacaccgc tccgtcctcc 3600
gacttcaata aggagcttac tcctccttga caccacccct ccagttcttc tcggcgctct 3660
ggagggaggc cttgtcggtc ttgggctggc cctggctgag aaagctgttg gcagccttaa 3720
agggagcgtg gaggtcacca gtcgctggct tcccgaagac gtggatctta accagattcg 3780
aaagcgcctt cagcggtatg tcgactggat cagaagagcg ttggtgtact tgaagtacag 3840
atgcatgacg gccatcatgc caacgcocat gaactggctc ttaatgagct ggcggaactg 3900
gcccttatcg tactccatgt tggtagttgt gacaggacga ggctcctcgc cgcttccaag 3960
cggagcaggc tcgacgtatt tcagtgtcga aagatctgat caagagacag gatgaggatc 4020
gtttcgcatt attgaacaag atggattgca cgcaggttct ccggccgctt ggggtggagag 4080
gctattcggc tatgactggg cacaacagac aatcggtgc tctgatgccg ccgtgttccg 4140
gctgtcagcg caggggcgcc cggttctttt tgtcaagacc gacctgtccg gtgcctgaa 4200
tgaactgcag gacgaggcag cgcggctatc gtggctggcc acgacgggag ttccttgccg 4260
agctgtgctc gacgttgtca ctgaagcggg aagggaactg ctgctatttg gcgaagtgcc 4320
ggggcaggat ctctgtcat ctcaccttgc tcctgccgag aaagtatcca tcatggctga 4380

tgcaatgcgg cggtgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa 4440
acatcgcatc gagcgagcac gtactcggat ggaagccggt cttgtcgatc aggatgatct 4500
ggacgaagag catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgcgcat 4560
gcccgcggc gaggatctcg tcgtgaccca tggcgatgcc tgcttgccga atatcatggt 4620
ggaaaatggc cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta 4680
tcaggacata gcgttggtta cccgtgatat tgctgaagag cttggcggcg aatgggctga 4740
ccgcttcctc gtgctttacg gtatcgccgc tcccgattcg cagcgcatcg ccttctatcg 4800
ccttcttgac gagttcttct gagcgggact ctggggttcg aaatgaccga ccaagcgacg 4860
cccaacctgc catcacgaga ttctgattcc accgcgcct tctatgaaag gttgggcttc 4920
ggaatcgttt tccgggacgc cggctggatg atcctccagc gcggggatct catgctggag 4980
ttcttcgcc accccgggct cgatccctc gcgagttggt tcagctgctg cctgaggctg 5040
gacgacctcg cggagttcta ccggcagtgc aaatccgtcg gcatccagga aaccagcagc 5100
ggctatccgc gcatccatgc cccgaactg caggagtggg gaggcacgat ggccgctttg 5160
gtccggatct ttgtgaagga accttacttc tgtggtgtga cataattgga caaactacct 5220
acagagattt aaagctctaa ggtaaataa aaatttttaa gtgtataatg tgttaacta 5280
ctgattctaa ttgtttgtgt attttagatt ccaacctatg gaactgatga atgggagcag 5340
tggtggaatg cctttaatga ggaaaacctg ttttgctcag aagaaatgcc atctagtgat 5400
gatgaggcta ctgctgactc tcaacattct actcctcaa aaaagaagag aaaggtagaa 5460
gacccaagg actttccttc agaattgcta agttttttga gtcattgctgt gtttagtaat 5520
agaactcttg cttgctttgc tatttacacc acaaaggaaa aagctgcact gctatacaag 5580
aaaattatgg aaaaatattc tgtaaccttt ataagtaggc ataacagtta taatcataac 5640

atactgtttt ttcttactcc acacagggcat agagtgtctg ctattaataa ctatgctcaa 5700

aaattgtgta ccttttagctt ttttaatttgt aaaggggtta ataaggaata tttgatgtat 5760

agtgccctga ctagagatca taatcagcca taccacattt gtagaggttt tacttgcttt 5820

aaaaaacctc ccacacctcc ccttgaacct gaaacataaa atgaatgcaa ttgttggtgt 5880

taacttggtt attgcagctt ataatgggtta caaataaagc aatagcatca caaatttcac 5940

aaataaagca tttttttcac tgcattctag ttgtggtttg tccaaactca tcaatgtatc 6000

ttatcatgtc tggatctgac gggcgccat gatcgtgctc ctgctgttga ggacccggct 6060

aggctggcgg ggttgcctta ctggttagca gaatgaatca ccgatacgcg agcgaacgtg 6120

aagcgactgc tgctgcaaaa cgtctgcgac ctgagcaaca acatgaatgg tcttcggttt 6180

ccgtgtttcg taaagtctgg aaacgcggaa gtcagcgctc ttccgcttcc tcgctcactg 6240

actcgctgcg ctccggtcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa 6300

tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 6360

aaaaggccag caaaaggcca ggaaccgtta aaaggccgcg ttgctggcgt ttttccatag 6420

gctccgcccc cctgacgagc atcacaaaaa tcgacgtca agtcagaggt ggcgaaaccc 6480

gacaggacta taaagatacc aggcgtttcc ccttgaagc tccctcgtgc gctctcctgt 6540

tccgaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 6600

ttctcatagc tcacgtgta ggtatctcag ttcgggttag gtcgttcgct ccaagctggg 6660

ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc ttatccggtta actatcgtct 6720

tgagtccaac ccggttaagac acgacttatc gccactggca gcagccactg gtaacaggat 6780

tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtggc ctaactacgg 6840

ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa 6900
aagagttggg agctcttgat ccggcaaaca aaccaccgct ggtagcgggtg gtttttttgt 6960
ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 7020
tacgggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggatttttg tcatgagatt 7080
atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 7140
aagtatatat gagtaaacct ggtctgacag ttaccaatgc ttaatcagtg aggcacctat 7200
ctcagcgate tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac 7260
tacgatacgg gagggcttac catctggccc cagtgtctga atgataccgc gagaccacg 7320
ctcaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag 7380
tggtcctgca actttatccg cctccatcca gtctattaat tggtgccggg aagctagagt 7440
aagtagttcg ccagttaata gtttgcgcaa cggtgttgcc attgctgcag gcatcgtggg 7500
gtcacgctcg tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt 7560
tacatgatcc cccatgttgt gcaaaaaagc ggtagctcc ttcggctctc cgatcgttgt 7620
cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct 7680
tactgtcatg ccatccgtaa gatgcttttc tgtgactggg gagtactcaa ccaagtcatt 7740
ctgagaatag tgtatcggcg gaccgagttg ctcttgcccg gcgtcaacac gggataatac 7800
cgcgccacat agcagaactt taaaagtgt catcattgga aaacgttctt cggggcgaaa 7860
actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc gtgcacccaa 7920
ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca 7980
aaatgccgca aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttctt 8040
ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 8100

atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc 8160

tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggg gtatcacgag 8220

gccctttcgt cttcaagaat tcgcggccgc aattaaccct cactaaagga tccctatagt 8280

gagtcgtatt atgcggccgc gaattctcat gtttgaccgc ttatcatcga taagctctgc 8340

tttttggtga cttccattgt tcattccacg gacaaaaaca gagaaaggaa acgacagagg 8400

ccaaaaagct cgctttcagc acctgtcgtt tcctttcttt tcagagggta ttttaaataa 8460

aaacattaag ttatgacgaa gaagaacgga aacgccttaa accggaaaat tttcataaat 8520

agcgaaaacc cgcgaggtcg ccgccccgta acaaggcgga tcgccggaaa ggacccgcaa 8580

atgataataa ttatcaattg catactatcg acggcactgc tgccagataa caccaccggg 8640

gaaacattcc atcatgatgg ccgtgcggac ataggaagcc agttcatcca tcgctttctt 8700

gtctgctgcc atttgctttg tgacatccag cgccgcacat tcagcagcgt ttttcagcgc 8760

gttttcgacg aacgtttcaa tgttggtatc aacaccagggt ttaactttga acttatoggc 8820

actgacggtt accttgttct gcgctggctc atcacgcagg ataccaaggc tgatgttgta 8880

gatattggtc accggctgag ggttttcgat tgccgctgcg tggatagcac catttgcgat 8940

caggcngtcc ttgatgaatg acactccatt gcgaataagt tcgaaggaga cgggtgcacg 9000

aatgcgctgg tccagctcgg tcgattgcct tttgtgcagc agaggatatca atctcaacgc 9060

caaggctcat cgaagcgcaa tattgctgct caccaaaacg cgtattgacc aggtgttcaa 9120

cggcaaatct ctgcccttct gatgtcagaa aggcaaagtg attttctttc tggatttcag 9180

ttgctgtgtg tcggtttcag caaaaccaag ctgcgcgaat tcggctgtgc agatttagaa 9240

ggcagatcac cagacagcaa cggccaacgg aaaacagcgc atacagaaca tccgtcgccg 9300

cgccgacaac gtgataatct ttatgacca tgatttatct ccttttagac gtgagcctgt 9360

cgcacagcaa agccgccgaa agttcctcga agctagcttc agacgtgtct agatacgtct 9420

gctttttgtt gacttccatt gttcattcca cggacaaaaa cagagaaagg aaacgacaga 9480

ggccaaaaag ctcgctttca gcacctgtcg tttcctttct tttcagaggg tattttaaat 9540

aaaaacatta agttatgacg aagaagaacg gaaacgcctt aaaccggaaa attttcataa 9600

atagcgaaaa cccgcgaggt cgccgccccg taacaaggcg gatcgccgga aaggacccgc 9660

aatgataat aattatcaat tgcatactat cgacggcact gctgccagat aacaccaccg 9720

gggaaacatt ccatcatgat ggccgtgcgg acataggaag ccagttcatc catcgctttc 9780

ttgtctgctg ccatttgctt tgtgacatcc agcgcgcac attcagcagc gtttttcagc 9840

gcgttttoga tcaacgtttc aatgttgga tcaacaccag gtttaacttt gaacttatcg 9900

gactgacgg ttaccttggt ctgcgtggc tcatcacgca ggataccaag gctgatgttg 9960

tagatatggg tcaccggctg aggggtttcg attgccgctg cgtggatagc accatttgcg 10020

atcaggcngt ccttgatgaa tgacactcca ttgcgaataa gttcgaagga gacggtgtca 10080

cgaatgcgct ggtccagctc ggtcgattgc cttttgtgca gcagaggat caatctcaac 10140

gccaaaggctc atcgaagcgc aatattgctg ctcacaaaaa cgcgtattga ccagggtgttc 10200

aacggcaaat ttctgccctt ctgatgtcag aaaggcaaag tgattttctt tctggtattc 10260

agttgctgtg tgtcggtttc agcaaaacca agctcgcgca attcggctgt gcagatttag 10320

aaggcagatc accagacagc aacggccaac ggaaaacagc gcatacagaa catccgtcgc 10380

cgcgcgcgaca acgtgataat ttttatgacc catgatttat ttcccttttag acgtgagcct 10440

gtcgcacagc aaagccgccg aaagttcctc gaccgatgcc cttgagagcc ttcaaccag 10500

tcagctcctt ccggtgggcg cggggcatga ctatcgtcgc cgcacttatg actgtcttct 10560

ttatcatgca actcgtagga cagggtgccg cagcgtctg ggtcattttc ggcgaggacc 10620

gctttcgctg gagcgcgacg atgacgagcc tgctgcttgc ggtattcgga atcttgcacg 10680
ccctcgctca agccttcgtc actgggtccg ccaccaaagc ttctggcgag aagcaggcca 10740
ttatcgccgg catggcgccg gacgcgctgg gctacgtctt gctggcgctc gcgacgcgag 10800
gctggatggc cttccccatt atgattcttc tcgcttccgg cggcatcggg atgcccgcgt 10860
tgcaggccat gctgtccagg caggtagatg acgaccatca gggacagctt caaggatcgc 10920
tcgcggtctt taccagccta acttcgatca ttggaccgct gatcgtcacg gcgatttatg 10980
ccgcctcggc gagcacatgg aacgggttgg catggattgt aggcgcccgc ctataccttg 11040
tctgcctccc cgcgttgcgt cgcggtgcat ggagccgggc cacctcgacc tgaatggaag 11100
ccggcggcac ctgcctaagc gattcaccac tccaagaatt ggagccaatc aattcttgcg 11160
gagaactgtg aatgcgcaaa ccaacccttg gcagaacata tccatcgctt ccgccatctc 11220
cagcagccgc acgcggcgca tctcgggcag cgttgggtcc tgcagatccg gctgtggaat 11280
gtgtgtcagt tagggtgtgg aaagtcccca ggctccccag caggcagaag tatgcaaagc 11340
atgcatctca attagtcagc aaccagggtg ggaaagtccc caggctcccc agcaggcaga 11400
agtatgcaaa gcatgcatct caattagtca gcaaccatag tcccgcccct aactccgccc 11460
atcccgcccc taactccgcc cagttccgcc cattctccgc cccatggctg actaattttt 11520
tttatattatg cagaggccga ggccgcctcg gcctctgagc tattccagaa gtagtgagga 11580
ggcttttttg gaggcctagg cttttgcaaa a 11611

<210> 5

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 5

cgatgtagga gggcgtggat a

21

<210> 6

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 6

gcttctgcgg gcgatttg t

21

<210> 7

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 7

tgagaatatc accggaattg

20

<210> 8

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 8

agctogacat actgttcttc c

21

9

24

, > DNA

13> Artificial

<220>

<223> Primer

<400> 9

gtgaatggaa atcccatcgc tgtc

24

<210> 10

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 10

agtgggtact ctaaaggcca tacc

24

<210> 11

<211> 1771

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (166)..(1155)

<400> 11

ggcacgagct tgcaagcaag tcagcgcgcg caagtcaaca cctgccggtc cacagcctca . 60

aataataaag agtcaagcg tttgtgcgcc tcgacgtggc cagtctgcac tgccttgaac 120

ccgcgagtct cccgccgcac tgactgccat agcacagcta gacga atg cag cta gca 177

Met Gln Leu Ala

gcg aca gta atg ttg gag cag ctt acc gga agc gct gag gca ctc aag 225
 Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala Glu Ala Leu Lys
 5 10 15 20

gag aag gag aag gag gtt gca ggc agc tct gac gtg ttg cgt aca tgg 273
 Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val Leu Arg Thr Trp
 25 30 35

gcg acc cag tac tcg ctt ccg tca gaa gag tca gac gcg gcc cgc ccg 321
 Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp Ala Ala Arg Pro
 40 45 50

gga ctg aag aat gcc tac aag cca cca cct tcc gac aca aag ggc atc 369
 Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp Thr Lys Gly Ile
 55 60 65

aca atg gcg cta cgt gtc atc ggc tcc tgg gcc gca gtg ttc ctc cac 417
 Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala Val Phe Leu His
 70 75 80

gcc att ttt caa atc aag ctt ccg acc tcc ttg gac cag ctg cac tgg 465
 Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp Gln Leu His Trp
 85 90 95 100

ctg ccc gtg tca gat gcc aca gct cag ctg gtt agc ggc acg agc agc 513
 Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser Gly Thr Ser Ser
 105 110 115

ctg ctc gac atc gtc gta gta ttc ttt gtc ctg gag ttc ctg tac aca 561
 Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu Phe Leu Tyr Thr
 120 125 130

ggc ctt ttt atc acc acg cat gat gct atg cat ggc acc atc gcc atg 609
 Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly Thr Ile Ala Met
 135 140 145

aga aac agg cag ctt aat gac ttc ttg ggc aga gta tgc atc tcc ttg 657
 Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val Cys Ile Ser Leu
 150 155 160

tac gcc tgg ttt gat tac aac atg ctg cac cgc aag cat tgg gag cac	705
Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys His Trp Glu His	
165 170 175 180	
cac aac cac act ggc gag gtg ggc aag gac cct gac ttc cac agg gga	753
His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp Phe His Arg Gly	
185 190 195	
aac cct ggc att gtg ccc tgg ttt gcc agc ttc atg tcc agc tac atg	801
Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met	
200 205 210	
tcg atg tgg cag ttt gcg cgc ctc gca tgg tgg acg gtg gtc atg cag	849
Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr Val Val Met Gln	
215 220 225	
ctg ctg ggt gcg cca atg gcg aac ctg ctg gtg ttc atg gcg gcc gcg	897
Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala	
230 235 240	
ccc atc ctg tcc gcc ttc cgc ttg ttc tac ttt ggc acg tac atg ccc	945
Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Met Pro	
245 250 255 260	
cac aag cct gag cct ggc gcc gcg tca ggc tct tca cca gcc gtc atg	993
His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser Pro Ala Val Met	
265 270 275	
aac tgg tgg aag tcg cgc act agc cag gcg tcc gac ctg gtc agc ttt	1041
Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp Leu Val Ser Phe	
280 285 290	
ctg acc tgc tac cac ttc gac ctg cac tgg gag cac cac cgc tgg ccc	1089
Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His His Arg Trp Pro	
295 300 305	
ttc gcc ccc tgg tgg gag ctg ccc aac tgc cgc cgc ctg tct ggc cga	1137
Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg Leu Ser Gly Arg	
310 315 320	
ggt ctg gtt cct gcc tag ctggacacac tgcagtgggc cctgctgcc	1185
Gly Leu Val Pro Ala	

325

gctgggcatg caggttgtgg caggactggg tgaggtgaaa agctgcaggc gctgctgccg 1245
 gacacgctgc atgggctacc ctgtgtagct gccgccacta ggggagggggg tttgtagctg 1305
 tcgagcttgc cccatggatg aagctgtgta gtggtgcagg gactacaccc acaggccaac 1365
 acccttgcag gagatgtctt gcgtcgggag gagtgttggg cagtgtagat gctatgattg 1425
 tatcttaatg ctgaagcctt taggggagcg acacttagtg ctgggcaggc aacgcctgc 1485
 aaggtgcagg cacaagctag gctggacgag gactcggtag caggcagggtg aagagggtgcg 1545
 ggaggggtgg gccacacca ctgggcaaga ccatgctgca atgctggcgg tgtggcagtg 1605
 agagctgcgt gattaactgg gctatggatt gtttgagcag tctcacttat tctttgatat 1665
 agatactggg caggcaggtc aggagagtga gtatgaacaa gttgagaggt ggtgcgctgc 1725
 ccctgcgctt atgaagctgt aacaataaag tggttcaaaa aaaaaa 1771

<210> 12

<211> 329

<212> PRT

<213> Haematococcus pluvialis

<400> 12

Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala
 1 5 10 15

Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val
 20 25 30

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp
 35 40 45

33/357

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp
50 55 60

Thr Lys Gly Ile Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala
65 70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp
85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser
100 105 110

Gly Thr Ser Ser Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu
115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly
130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val
145 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys
165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp
180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met
195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr

210 215 220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe
225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly
245 250 255

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser
260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp
275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His
290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg
305 310 315 320

Leu Ser Gly Arg Gly Leu Val Pro Ala
325

<210> 13

<211> 1662

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (168) .. (1130)

<400> 13

cggggcgaact caagaaattc aacagctgca agcgcgcccc agcctcacag cgccaagtga 60
 gctatcgacg tggttgtgag cgctcgacgt ggtccactga cgggcctgtg agcctctgag 120
 ctccgtcctc tgccaaatct cgcgtcgggg cctgcctaag tcgaaga atg cac gtc 176
 Met His Val
 1
 gca tcg gca cta atg gtc gag cag aaa ggc agt gag gca gct gct tcc 224
 Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala Ala Ala Ser
 5 10 15
 agc cca gac gtc ttg aga gcg tgg gcg aca cag tat cac atg cca tcc 272
 Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His Met Pro Ser
 20 25 30 35
 gag tcg tca gac gca gct cgt cct gcg cta aag cac gcc tac aaa cct 320
 Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala Tyr Lys Pro
 40 45 50
 cca gca tct gac gcc aag ggc atc acg atg gcg ctg acc atc att ggc 368
 Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr Ile Ile Gly
 55 60 65
 acc tgg acc gca gtg ttt tta cac gca ata ttt caa atc agg cta ccg 416
 Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile Arg Leu Pro
 70 75 80
 aca tcc atg gac cag ctt cac tgg ttg cct gtg tcc gaa gcc aca gcc 464
 Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu Ala Thr Ala
 85 90 95
 cag ctt ttg ggc gga agc agc agc cta ctg cac atc gct gca gtc ttc 512
 Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala Ala Val Phe
 100 105 110 115
 att gta ctt gag ttc ctg tac act ggt cta ttc atc acc aca cat gac 560
 Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp
 120 125 130
 gca atg cat ggc acc ata gct ttg agg cac agg cag ctc aat gat ctc 608
 Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu Asn Asp Leu

135	140	145	
ctt ggc aac atc tgc ata tca ctg tac gcc tgg ttt gac tac agc atg			656
Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Ser Met			
150	155	160	
ctg cat cgc aag cac tgg gag cac cac aac cat act ggc gaa gtg ggg			704
Leu His Arg Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly			
165	170	175	
aaa gac cct gac ttc cac aag gga aat ccc ggc ctt gtc ccc tgg ttc			752
Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe			
180	185	190	195
gcc agc ttc atg tcc agc tac atg tcc ctg tgg cag ttt gcc cgg ctg			800
Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu			
200	205	210	
gca tgg tgg gca gtg gtg atg caa atg ctg ggg gcg ccc atg gca aat			848
Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro Met Ala Asn			
215	220	225	
ctc cta gtc ttc atg gct gca gcc cca atc ttg tca gca ttc cgc ctc			896
Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu			
230	235	240	
ttc tac ttc ggc act tac ctg cca cac aag cct gag cca ggc cct gca			944
Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala			
245	250	255	
gca ggc tct cag gtg atg gcc tgg ttc agg gcc aag aca agt gag gca			992
Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr Ser Glu Ala			
260	265	270	275
tct gat gtg atg agt ttc ctg aca tgc tac cac ttt gac ctg cac tgg			1040
Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp			
280	285	290	
gag cac cac agg tgg ccc ttt gcc ccc tgg tgg cag ctg ccc cac tgc			1088
Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu Pro His Cys			
295	300	305	

cgc cgc ctg tcc ggg cgt ggc ctg gtg cct gcc ttg gca tga 1130
 Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala
 310 315 320

cctgggtccct ccgctggtga cccagcgtct gcacaagagt gtcattgctac aggggtgctgc 1190
 ggccagtggc agcgagtgct actctcagcc tgtatggggc taccgctgtg cactgagca 1250
 ctgggcatgc cactgagcac tgggcgtgct actgagcaat gggcgtgcta ctgagcaatg 1310
 ggcgtgctac tgacaatggg cgtgctactg gggctctggc gtggctagga tggagtttga 1370
 tgcattcagt agcgggtggc aacgtcatgt ggatggtgga agtgctgagg ggtttaggca 1430
 gccggcattt gagagggcta agttataaat cgcattgctgc tcatgcgcac atatctgcac 1490
 acagccaggg aatcccttc gagagtgatt atgggacact tgtattgggt tcgtgctatt 1550
 gttttattca gcagcagtag ttagtgaggg tgagagcagg gtggtgagag tggagtgagt 1610
 gagtatgaac ctggtcagcg aggtgaacag cctgtaatga atgactctgt ct 1662

<210> 14

<211> 320

<212> PRT

<213> Haematococcus pluvialis

<400> 14

Met His Val Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala
 1 5 10 15

Ala Ala Ser Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His
 20 25 30

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala
 35 40 45

38/357

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr
50 55 60

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile
65 70 75 80

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu
85 90 95

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala
100 105 110

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr
115 120 125

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu
130 135 140

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp
145 150 155 160

Tyr Ser Met Leu His Arg Lys His Trp Glu His His Asn His Thr Gly
165 170 175

Glu Val Gly Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val
180 185 190

Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe
195 200 205

Ala Arg Leu Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro

210 215 220

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg 48
Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
1 5 10 15

atc gtc tcg ggc ggc atc atc gcc gct tgg ctg gcc ctg cat gtg cat	96
Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His	
20 25 30	
gcg ctg tgg ttt ctg gac gca gcg gcg cat ccc atc ctg gcg atc gca	144
Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala	
35 40 45	
aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg	192
Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala	
50 55 60	
cat gac gcg atg cac ggg tcg gtg gtg ccg ggg cgt ccg cgc gcc aat	240
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn	
65 70 75 80	
gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg	288
Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp	
85 90 95	
cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc	336
Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr	
100 105 110	
gac gac gac ccc gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc	384
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala	
115 120 125	
cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc	432
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro	
130 135 140	
gtc atc gtg acg gtc tat gcg ctg atc ctt ggg gat cgc tgg atg tac	480
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr	
145 150 155 160	
gtg gtc ttc tgg ccg ctg ccg tcg atc ctg gcg tcg atc cag ctg ttc	528
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe	
165 170 175	
gtg ttc ggc acc tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg	576
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro	

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp

225

230

235

240

Thr Ala

<210> 17

<211> 1631

<212> DNA

<213> Alcaligenes sp.

<220>

<221> CDS

<222> (99)..(827)

<400> 17

ctgcaggccg ggcccgggtgg ccaatggtcg caaccggcag gactggaaca ggacggcggg 60

ccggctctagg ctgtcgccct acgcagcagg agtttcgg atg tcc gga cgg aag cct 116

Met Ser Gly Arg Lys Pro

1

5

ggc aca act ggc gac acg atc gtc aat ctc ggt ctg acc gcc gcg atc 164

Gly Thr Thr Gly Asp Thr Ile Val Asn Leu Gly Leu Thr Ala Ala Ile

10

15

20

ctg ctg tgc tgg ctg gtc ctg cac gcc ttt acg cta tgg ttg cta gat 212

Leu Leu Cys Trp Leu Val Leu His Ala Phe Thr Leu Trp Leu Leu Asp

25

30

35

gcg gcc gcg cat ccg ctg ctt gcc gtg ctg tgc ctg gct ggg ctg acc 260

Ala Ala Ala His Pro Leu Leu Ala Val Leu Cys Leu Ala Gly Leu Thr

40

45

50

tgg ctg tgc gtc ggg ctg ttc atc atc gcg cat gac gca atg cac ggg 308

Trp Leu Ser Val Gly Leu Phe Ile Ile Ala His Asp Ala Met His Gly

55

60

65

70

tcc gtg gtg ccg ggg cgg ccg cgc gcc aat gcg gcg atc ggg caa ctg 356

Ser Val Val Pro Gly Arg Pro Arg Ala Asn Ala Ala Ile Gly Gln Leu

75	80	85	
gcg ctg tgg ctc tat gcg ggg ttc tcg tgg ccc aag ctg atc gcc aag			404
Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp Pro Lys Leu Ile Ala Lys			
90	95	100	
cac atg acg cat cac cgg cac gcc ggc acc gac aac gat ccc gat ttc			452
His Met Thr His His Arg His Ala Gly Thr Asp Asn Asp Pro Asp Phe			
105	110	115	
ggt cac gga ggg ccc gtg cgc tgg tac ggc agc ttc gtc tcc acc tat			500
Gly His Gly Gly Pro Val Arg Trp Tyr Gly Ser Phe Val Ser Thr Tyr			
120	125	130	
ttc ggc tgg cga gag gga ctg ctg cta ccg gtg atc gtc acc acc tat			548
Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro Val Ile Val Thr Thr Tyr			
135	140	145	150
gcg ctg atc ctg ggc gat cgc tgg atg tat gtc atc ttc tgg ccg gtc			596
Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr Val Ile Phe Trp Pro Val			
155	160	165	
ccg gcc gtt ctg gcg tcg atc cag att ttc gtc ttc gga act tgg ctg			644
Pro Ala Val Leu Ala Ser Ile Gln Ile Phe Val Phe Gly Thr Trp Leu			
170	175	180	
ccc cac cgc ccg gga cat gac gat ttt ccc gac ccg cac aac gcg agg			692
Pro His Arg Pro Gly His Asp Asp Phe Pro Asp Arg His Asn Ala Arg			
185	190	195	
tcg acc ggc atc ggc gac ccg ttg tca cta ctg acc tgc ttc cat ttc			740
Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu Leu Thr Cys Phe His Phe			
200	205	210	
ggc ggc tat cac cac gaa cat cac ctg cat ccg cat gtg ccg tgg tgg			788
Gly Gly Tyr His His Glu His His Leu His Pro His Val Pro Trp Trp			
215	220	225	230
cgc ctg cct cgt aca cgc aag acc gga ggc cgc gca tga cgcaattcct			837
Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly Arg Ala			
235	240		

cattgtcgtg gcgacagtc tctgatgga gctgaccgcc tattccgtcc accgctggat 897
 tatgcacggc ccctaggct ggggctggca caagtcccat cacgaagagc acgaccacgc 957
 gttggagaag aacgacctct acggcgtcgt cttegcggtg ctggcgacga tcctcttcac 1017
 cgtgggggcc tattggtggc cggctgctgt gtggatcgcc ctgggcatga cggctctatgg 1077
 gttgatctat ttcactctgc acgacgggct tgtgcatcaa cgctggccgt ttcggtatat 1137
 tccggggcgg ggctatttcc gcaggctcta ccaagctcat cgcctgcacc acgcggtcga 1197
 ggggcgggac cactgcgtca gcttcggctt catctatgcc ccaccgtgg acaagctgaa 1257
 gcaggatctg aagcggtcgg gtgtcctgcg ccccaggac gagcgtccgt cgtgatctct 1317
 gatcccggcg tggccgcatg aaatccgacg tctgtctggc aggggcccgc cttgccaacg 1377
 gactgatcgc gctggcgatc cgcaaggcgc ggcccagcct tcgcgtgctg ctgctggacc 1437
 gtgcggcggg cgcctcggac gggcatactt ggtcctgcc cgacaccgat ttggcgccgc 1497
 actggctgga ccgcctgaag ccgatcaggc gtggcgactg gcccgatcag gaggtgcggt 1557
 tcccagacca ttcgcgaagg ctccgggccg gatatggctc gatcgacggg cgggggctga 1617
 tgcgtgcggt gacc 1631

<210> 18

<211> 242

<212> PRT

<213> *Alcaligenes* sp.

<400> 18

Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu
 1 5 10 15

Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe

20	25	30
Thr Leu Trp Leu Leu Asp Ala Ala Ala His Pro Leu Leu Ala Val Leu		
35	40	45
Cys Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala		
50	55	60
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn		
65	70	75
Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp		
85	90	95
Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr		
100	105	110
Asp Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly		
115	120	125
Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro		
130	135	140
Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr		
145	150	155
Val Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe		
165	170	175
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro		
180	185	190

Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu
 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
 210 215 220

Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly
 225 230 235 240

Arg Ala

<210> 19

<211> 729

<212> DNA

<213> *Paracoccus marcusii*

<220>

<221> CDS

<222> (1)..(729)

<400> 19

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc aca agc ctg 48

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu

1 5 10 15

atc gtc tcg ggc ggc atc atc gcc gca tgg ctg gcc ctg cat gtg cat 96

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His

20 25 30

gcg ctg tgg ttt ctg gac gcg gcg gcc cat ccc atc ctg gcg gtc gcg 144

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala

35 40 45

aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg 192

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala

50	55	60	
cat gac gcg atg cac ggg tcg gtc gtg ccg ggg cgt ccg cgc gcc aat			240
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn			
65	70	75	80
gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg			288
Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp			
85	90	95	
cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc			336
Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr			
100	105	110	
gac gac gac cca gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc			384
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala			
115	120	125	
cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc			432
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro			
130	135	140	
gtc atc gtg acg gtc tat gcg ctg atc ctg ggg gat cgc tgg atg tac			480
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr			
145	150	155	160
gtg gtc ttc tgg ccg ttg ccg tcg atc ctg gcg tcg atc cag ctg ttc			528
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe			
165	170	175	
gtg ttc ggc act tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg			576
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro			
180	185	190	
gac cgc cat aat gcg cgg tcg tcg ccg atc agc gac cct gtg tcg ctg			624
Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu			
195	200	205	
ctg acc tgc ttt cat ttt ggc ggt tat cat cac gaa cac cac ctg cac			672
Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His			
210	215	220	

ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac 720
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

acc gca tga 729
Thr Ala

<210> 20

<211> 242

<212> PRT

<213> Paracoccus marcusii

<400> 20

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala
35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

Thr Ala

<210> 21

<211> 1629

<212> DNA

<213> Synechocystis sp.

<220>

<221> CDS

<222> (1)..(1629)

<400> 21

atg atc acc acc gat gtt gtc att att ggg gcg ggg cac aat ggc tta	48
Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu	
1 5 10 15	
gtc tgt gca gcc tat ttg ctc caa cgg ggc ttg ggg gtg acg tta cta	96
Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu	
20 25 30	
gaa aag cgg gaa gta cca ggg ggg gcg gcc acc aca gaa gct ctc atg	144
Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met	
35 40 45	
ccg gag cta tcc ccc cag ttt cgc ttt aac cgc tgt gcc att gac cac	192
Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His	
50 55 60	
gaa ttt atc ttt ctg ggg ccg gtg ttg cag gag cta aat tta gcc cag	240
Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln	
65 70 75 80	
tat ggt ttg gaa tat tta ttt tgt gac ccc agt gtt ttt tgt ccg ggg	288
Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly	
85 90 95	
ctg gat ggc caa gct ttt atg agc tac cgt tcc cta gaa aaa acc tgt	336
Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys	
100 105 110	
gcc cac att gcc acc tat agc ccc cga gat gcg gaa aaa tat cgg caa	384
Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln	
115 120 125	
ttt gtc aat tat tgg acg gat ttg ctc aac gct gtc cag cct gct ttt	432

290	295	300	
caa ttg gtg gaa ccg ggg gcc cta gcc aag gtg aat caa aac cta ggg			960
Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly			
305	310	315	320
gaa cga ctg gaa cgg cgc act gtg aac aat aac gaa gcc att tta aaa			1008
Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys			
325	330	335	
atc gat tgt gcc ctc tcc ggt tta ccc cac ttc act gcc atg gcc ggg			1056
Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly			
340	345	350	
ccg gag gat cta acg gga act att ttg att gcc gac tcg gta cgc cat			1104
Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His			
355	360	365	
gtc gag gaa gcc cac gcc ctc att gcc ttg ggg caa att ccc gat gct			1152
Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala			
370	375	380	
aat ccg tct tta tat ttg gat att ccc act gta ttg gac ccc acc atg			1200
Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met			
385	390	395	400
gcc ccc cct ggg cag cac acc ctc tgg atc gaa ttt ttt gcc ccc tac			1248
Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr			
405	410	415	
cgc atc gcc ggg ttg gaa ggg aca ggg tta atg ggc aca ggt tgg acc			1296
Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr			
420	425	430	
gat gag tta aag gaa aaa gtg gcg gat cgg gtg att gat aaa tta acg			1344
Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr			
435	440	445	
gac tat gcc cct aac cta aaa tct ctg atc att ggt cgc cga gtg gaa			1392
Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu			
450	455	460	

agt ccc gcc gaa ctg gcc caa cgg ctg gga agt tac aac ggc aat gtc 1440
 Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val
 465 470 475 480

tat cat ctg gat atg agt ttg gac caa atg atg ttc ctc cgg cct cta 1488
 Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
 485 490 495

ccg gaa att gcc aac tac caa acc ccc atc aaa aat ctt tac tta aca 1536
 Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
 500 505 510

ggg gcg ggt acc cat ccc ggt ggc tcc ata tca ggt atg ccc ggt aga 1584
 Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
 515 520 525

aat tgc gct cgg gtc ttt tta aaa caa caa cgt cgt ttt tgg taa 1629
 Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
 530 535 540

<210> 22

<211> 542

<212> PRT

<213> Synechocystis sp.

<400> 22

Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu
 1 5 10 15

Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu
 20 25 30

Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met

35

40

45

Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His
50 55 60

Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln
65 70 75 80

Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly
85 90 95

Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys
100 105 110

Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln
115 120 125

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe
130 135 140

Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp
145 150 155 160

Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala
165 170 175

Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn
180 185 190

Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys
195 200 205

Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met

210 215 220

Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly
225 230 235 240

Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln
245 250 255

Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu
260 265 270

Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg
275 280 285

Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu
290 295 300

Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly
305 310 315 320

Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys
325 330 335

Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly
340 345 350

Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His
355 360 365

Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala
370 375 380

Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met
385 390 395 400

Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr
405 410 415

Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr
420 425 430

Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr
435 440 445

Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu
450 455 460

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val
465 470 475 480

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
485 490 495

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
500 505 510

Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
515 520 525

Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
530 535 540

<211> 776

<212> DNA

<213> Bradyrhizobium sp.

<220>

<221> CDS

<222> (1)..(774)

<400> 23

atg cat gca gca acc gcc aag gct act gag ttc ggg gcc tct cgg cgc	48
Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg	
1 5 10 15	

gac gat gcg agg cag cgc cgc gtc ggt ctc acg ctg gcc gcg gtc atc	96
Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile	
20 25 30	

atc gcc gcc tgg ctg gtg ctg cat gtc ggt ctg atg ttc ttc tgg ccg	144
Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro	
35 40 45	

ctg acc ctt cac agc ctg ctg ccg gct ttg cct ctg gtg gtg ctg cag	192
Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln	
50 55 60	

acc tgg ctc tat gta ggc ctg ttc atc atc gcg cat gac tgc atg cac	240
Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His	
65 70 75 80	

ggc tcg ctg gtg ccg ttc aag ccg cag gtc aac cgc cgt atc gga cag	288
Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln	
85 90 95	

ctc tgc ctg ttc ctc tat gcc ggg ttc tcc ttc gac gct ctc aat gtc	336
Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val	
100 105 110	

gag cac cac aag cat cac cgc cat ccc ggc acg gcc gag gat ccc gat	384
Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp	
115 120 125	

ttc gac gag gtg ccg ccg cac ggc ttc tgg cac tgg ttc gcc agc ttt 432
 Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe
 130 135 140

ttc ctg cac tat ttc ggc tgg aag cag gtc gcg atc atc gca gcc gtc 480
 Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val
 145 150 155 160

tcg ctg gtt tat cag ctc gtc ttc gcc gtt ccc ttg cag aac atc ctg 528
 Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu
 165 170 175

ctg ttc tgg gcg ctg ccc ggg ctg ctg tcg gcg ctg cag ctg ttc acc 576
 Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
 180 185 190

ttc ggc acc tat ctg ccg cac aag ccg gcc acg cag ccc ttc gcc gat 624
 Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp
 195 200 205

cgc cac aac gcg cgg acg agc gaa ttt ccc gcg tgg ctg tcg ctg ctg 672
 Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu
 210 215 220

acc tgc ttc cac ttc ggc ttt cat cac gag cat cat ctg cat ccc gat 720
 Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp
 225 230 235 240

gcg ccg tgg tgg cgg ctg ccg gag atc aag cgg cgg gcc ctg gaa agg 768
 Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
 245 250 255

cgt gac ta 776
 Arg Asp

<210> 24

<211> 258

<212> PRT

<213> Bradyrhizobium sp.

<400> 24

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg
1 5 10 15

Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile
20 25 30

Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro
35 40 45

Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln
50 55 60

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His
65 70 75 80

Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln
85 90 95

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val
100 105 110

Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp
115 120 125

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe
130 135 140

Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val
145 150 155 160

61/357

Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu
 165 170 175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
 180 185 190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp
 195 200 205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu
 210 215 220

Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp
 225 230 235 240

Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
 245 250 255

Arg Asp

<210> 25

<211> 777

<212> DNA

<213> Nostoc sp.

<220>

<221> CDS

<222> (1) .. (777)

<400> 25

atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta
 Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu

1

5

10

15

ttg tca tcg aca atc aga gat gat aaa.aat att aat aag ggt ata ttt	96
Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe	
20 25 30	
att gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta	144
Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu	
35 40 45	
ctc tca ata gat aca tcc ata att cat aag agc tta tta ggt ata gcc	192
Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala	
50 55 60	
atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat	240
Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His	
65 70 75 80	
gat gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat	288
Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn	
85 90 95	
ttt ata ggt aag ctc act cta atc ttg tat gga cta ctc cct tat aaa	336
Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys	
100 105 110	
gat tta ttg aaa aaa cat tgg tta cac cac gga cat cct ggt act gat	384
Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp	
115 120 125	
tta gac cct gat tat tac aat ggt cat ccc caa aac ttc ttt ctt tgg	432
Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp	
130 135 140	
tat cta cat ttt atg aag tct tat tgg cga tgg acg caa att ttc gga	480
Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly	
145 150 155 160	
tta gtg atg att ttt cat gga ctt aaa aat ctg gtg cat ata cca gaa	528
Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu	
165 170 175	
aat aat tta att ata ttt tgg atg ata cct tct att tta agt tca gta	576

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val
 180 185 190

caa cta ttt tat ttt ggt aca ttt ttg cct cat aaa aag cta gaa ggt 624
 Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
 195 200 205

ggt tat act aac ccc cat tgt gcg cgc agt atc cca tta cct ctt ttt 672
 Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
 210 215 220

tgg tct ttt gtt act tgt tat cac ttc ggc tac cac aag gaa cat cac 720
 Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
 225 230 235 240

gaa tac cct caa ctt cct tgg tgg aaa tta cct gaa gct cac aaa ata 768
 Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
 245 250 255

tct tta taa 777
 Ser Leu

<210> 26
 <211> 258
 <212> PRT
 <213> Nostoc sp.

<400> 26

Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu
 1 5 10 15

Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe
 20 25 30

Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu
 35 40 45

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala
50 55 60

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His
65 70 75 80

Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn
85 90 95

Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys
100 105 110

Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp
115 120 125

Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp
130 135 140

Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly
145 150 155 160

Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu
165 170 175

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val
180 185 190

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
195 200 205

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
 210 215 220

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
 225 230 235 240

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
 245 250 255

Ser Leu

<210> 27

<211> 789

<212> DNA

<213> Nostoc punctiforme

<220>

<221> CDS

<222> (1)..(789)

<400> 27

ttg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa 48
 Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
 1 5 10 15

tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta 96
 Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
 20 25 30

att att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat 144
 Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
 35 40 45

tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa 192
 Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
 50 55 60

atg ttc ctt tat aca ggg cta ttt att act gca cat gat gct atg cat	240
Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His	
65 70 75 80	
ggg tca gtt tat cgt aaa aat ccc aaa att aat aat ttt atc ggt tca	288
Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser	
85 90 95	
cta gct gta gcg ctt tac gct gtg ttt cca tat caa cag atg tta aag	336
Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys	
100 105 110	
aat cat tgc tta cat cat cgt cat cct gct agc gaa gtt gac cca gat	384
Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp	
115 120 125	
ttt cat gat ggt aag aga aca aac gct att ttc tgg tat ctc cat ttc	432
Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe	
130 135 140	
atg ata gaa tac tcc agt tgg caa cag tta ata gta cta act atc cta	480
Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu	
145 150 155 160	
ttt aat tta gct aaa tac gtt ttg cac atc cat caa ata aat ctc atc	528
Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile	
165 170 175	
tta ttt tgg agt att cct cca att tta agt tcc att caa ctg ttt tat	576
Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr	
180 185 190	
ttc gga aca ttt ttg cct cat cga gaa ccc aag aaa gga tat gtt tat	624
Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr	
195 200 205	
ccc cat tgc agc caa aca ata aaa ttg cca act ttt ttg tca ttt atc	672
Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile	
210 215 220	
gct tgc tac cac ttt ggt tat cat gaa gaa cat cat gag tat ccc cat	720

67/357

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
225 230 235 240

gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac 768
Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
245 250 255

aat tca gta acc aat tcg taa 789
Asn Ser Val Thr Asn Ser
260

<210> 28

<211> 262

<212> PRT

<213> Nostoc punctiforme

<400> 28

Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
1 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
20 25 30

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
35 40 45

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
50 55 60

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His
65 70 75 80

Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser
85 90 95

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys
100 105 110

Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp
115 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe
130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu
145 150 155 160

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile
165 170 175

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr
180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr
195 200 205

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile
210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
225 230 235 240

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
245 250 255

Asn Ser Val Thr Asn Ser

260

<210> 29

<211> 762

<212> DNA

<213> Nostoc punctiforme

<220>

<221> CDS

<222> (1)..(762)

<400> 29

gtg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca 48

Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro

1 5 10 15

gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc 96

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val

20 25 30

att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac 144

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp

35 40 45

atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa 192

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln

50 55 60

aca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat 240

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His

65 70 75 80

ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca 288

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr

85 90 95

ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa 336

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys

100 105 110

aaa cat tgg tta cac cac cac aat cca gca agc tca ata gac ccg gat 384
Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp
115 120 125

ttt cac aat ggt aaa cac caa agt ttc ttt gct tgg tat ttt cat ttt 432
Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
130 135 140

atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att 480
Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
145 150 155 160

tat aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act 528
Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
165 170 175

tac ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat 576
Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
180 185 190

ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag 624
Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
195 200 205

cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc 672
Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
210 215 220

acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat 720
Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
225 230 235 240

att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag 762
Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
245 250

<210> 30

<211> 253

<212> PRT

<213> Nostoc punctiforme

<400> 30

Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro
1 5 10 15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val
20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp
35 40 45

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln
50 55 60

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His
65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr
85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys
100 105 110

Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp
115 120 125

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
130 135 140

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
 165 170 175

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
 180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
 195 200 205

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
 210 215 220

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
 225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
 245 250

<210> 31

<211> 1608

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (3) ..(971)

<400> 31

ct aca ttt cac aag ccc gtg agc ggt gca agc gct ctg ccc cac atc 47

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile

1 5 10 15

ggc cca cct cct cat ctc cat egg tca ttt gct gct acc acg atg ctg 95

Gly	Pro	Pro	Pro	His	Leu	His	Arg	Ser	Phe	Ala	Ala	Thr	Thr	Met	Leu	
				20					25					30		
tcg	aag	ctg	cag	tca	atc	agc	gtc	aag	gcc	cgc	cgc	gtt	gaa	cta	gcc	143
Ser	Lys	Leu	Gln	Ser	Ile	Ser	Val	Lys	Ala	Arg	Arg	Val	Glu	Leu	Ala	
			35					40					45			
cgc	gac	atc	acg	cgg	ccc	aaa	gtc	tgc	ctg	cat	gct	cag	cgg	tgc	tcg	191
Arg	Asp	Ile	Thr	Arg	Pro	Lys	Val	Cys	Leu	His	Ala	Gln	Arg	Cys	Ser	
			50					55					60			
tta	gtt	cgg	ctg	cga	gtg	gca	gca	cca	cag	aca	gag	gag	gcg	ctg	gga	239
Leu	Val	Arg	Leu	Arg	Val	Ala	Ala	Pro	Gln	Thr	Glu	Glu	Ala	Leu	Gly	
			65					70					75			
acc	gtg	cag	gct	gcc	ggc	gcg	ggc	gat	gag	cac	agc	gcc	gat	gta	gca	287
Thr	Val	Gln	Ala	Ala	Gly	Ala	Gly	Asp	Glu	His	Ser	Ala	Asp	Val	Ala	
80					85					90					95	
ctc	cag	cag	ctt	gac	cgg	gct	atc	gca	gag	cgt	cgt	gcc	cgg	cgc	aaa	335
Leu	Gln	Gln	Leu	Asp	Arg	Ala	Ile	Ala	Glu	Arg	Arg	Ala	Arg	Arg	Lys	
					100					105					110	
cgg	gag	cag	ctg	tca	tac	cag	gct	gcc	gcc	att	gca	gca	tca	att	ggc	383
Arg	Glu	Gln	Leu	Ser	Tyr	Gln	Ala	Ala	Ala	Ile	Ala	Ala	Ser	Ile	Gly	
					115					120					125	
gtg	tca	ggc	att	gcc	atc	ttc	gcc	acc	tac	ctg	aga	ttt	gcc	atg	cac	431
Val	Ser	Gly	Ile	Ala	Ile	Phe	Ala	Thr	Tyr	Leu	Arg	Phe	Ala	Met	His	
				130						135					140	
atg	acc	gtg	ggc	ggc	gca	gtg	cca	tgg	ggt	gaa	gtg	gct	ggc	act	ctc	479
Met	Thr	Val	Gly	Gly	Ala	Val	Pro	Trp	Gly	Glu	Val	Ala	Gly	Thr	Leu	
				145						150					155	
ctc	ttg	gtg	gtt	ggt	ggc	gcg	ctc	ggc	atg	gag	atg	tat	gcc	cgc	tat	527
Leu	Leu	Val	Val	Gly	Gly	Ala	Leu	Gly	Met	Glu	Met	Tyr	Ala	Arg	Tyr	
160						165						170			175	
gca	cac	aaa	gcc	atc	tgg	cat	gag	tcg	cct	ctg	ggc	tgg	ctg	ctg	cac	575
Ala	His	Lys	Ala	Ile	Trp	His	Glu	Ser	Pro	Leu	Gly	Trp	Leu	Leu	His	
					180							185			190	

aag agc cac cac aca cct cgc act gga ccc ttt gaa gcc aac gac ttg 623
Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu
195 200 205

ttt gca atc atc aat gga ctg ccc gcc atg ctc ctg tgt acc ttt ggc 671
Phe Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly
210 215 220

ttc tgg ctg ccc aac gtc ctg ggg gcg gcc tgc ttt gga gcg ggg ctg 719
Phe Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu
225 230 235

ggc atc acg cta tac ggc atg gca tat atg ttt gta cac gat ggc ctg 767
Gly Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu
240 245 250 255

gtg cac agg cgc ttt ccc acc ggg ccc atc gct ggc ctg ccc tac atg 815
Val His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met
260 265 270

aag cgc ctg aca gtg gcc cac cag cta cac cac agc ggc aag tac ggt 863
Lys Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly
275 280 285

ggc gcg ccc tgg ggt atg ttc ttg ggt cca cag gag ctg cag cac att 911
Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile
290 295 300

cca ggt gcg gcg gag gag gtg gag cga ctg gtc ctg gaa ctg gac tgg 959
Pro Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp
305 310 315

tcc aag cgg tag ggtgcggaac caggcacgct gggtttcacac ctcatgcctg 1011
Ser Lys Arg
320

tgataagggtg tggctagagc gatgcgtgtg agacgggtat gtcacgggtcg actggtctga 1071

tggccaatgg catcggccat gtctggtcat cacgggctgg ttgcctgggt gaaggtgatg 1131

cacatcatca tgtgcggttg gaggggctgg cacagtgtgg gctgaactgg agcagttgtc 1191

caggctggcg ttgaatcagt gagggtttgt gattggcggt tgtgaagcaa tgactccgcc 1251
catattctat ttgtgggagc tgagatgatg gcatgcttgg gatgtgcatg gatcatggta 1311
gtgcagcaaa ctatattcac ctagggctgt tggtaggata aggtgaggcc ttgcacattg 1371
catgatgtac tcgtcatggt gtgttgggtga gaggatggat gtggatggat gtgtattctc 1431
agacgtagac cttgactgga ggcttgatcg agagagtggg ccgtattctt tgagagggga 1491
ggctcgtgcc agaaatggtg agtggatgac tgtgacgctg tacattgcag gcaggtgaga 1551
tgcactgtct cgattgtaaa atacattcag atgcaaaaaa aaaaaaaaaa aaaaaaa 1608

<210> 32

<211> 322

<212> PRT

<213> Haematococcus pluvialis

<400> 32

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile Gly
1 5 10 15

Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu Ser
20 25 30

Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala Arg
35 40 45

Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser Leu
50 55 60

Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly Thr
65 70 75 80

Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala Leu
85 90 95

Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys Arg
100 105 110

Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly Val
115 120 125

Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His Met
130 135 140

Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu Leu
145 150 155 160

Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr Ala
165 170 175

His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His Lys
180 185 190

Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu Phe
195 200 205

Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly Phe
210 215 220

Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu Gly
225 230 235 240

Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu Val

245

250

255

His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met Lys
 260 265 270

Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly Gly
 275 280 285

Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile Pro
 290 295 300

Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp Ser
 305 310 315 320

Lys Arg

<210> 33

<211> 528

<212> DNA

<213> Erwinia uredovora

<220>

<221> CDS

<222> (1)..(528)

<400> 33

atg ttg tgg att tgg aat gcc ctg atc gtt ttc gtt acc gtg att ggc 48
 Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly
 1 5 10 15

atg gaa gtg att gct gca ctg gca cac aaa tac atc atg cac ggc tgg 96
 Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
 20 25 30

78/357

```

ggt tgg gga tgg cat ctt tca cat cat gaa ccg cgt aaa ggt gcg ttt      144
Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe
      35                      40                      45

gaa gtt aac gat ctt tat gcc gtg gtt ttt gct gca tta tcg atc ctg      192
Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu
      50                      55                      60

ctg att tat ctg ggc agt aca gga atg tgg ccg ctc cag tgg att ggc      240
Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly
      65                      70                      75                      80

gca ggt atg acg gcg tat gga tta ctc tat ttt atg gtg cac gac ggg      288
Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly
      85                      90                      95

ctg gtg cat caa cgt tgg cca ttc cgc tat att cca cgc aag ggc tac      336
Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr
      100                     105                     110

ctc aaa cgg ttg tat atg gcg cac cgt atg cat cac gcc gtc agg ggc      384
Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
      115                     120                     125

aaa gaa ggt tgt gtt tct ttt ggc ttc ctc tat gcg ccg ccc ctg tca      432
Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser
      130                     135                     140

aaa ctt cag gcg acg ctc cgg gaa aga cat ggc gct aga gcg ggc gct      480
Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala
      145                     150                     155                     160

gcc aga gat gcg cag ggc ggg gag gat gag ccc gca tcc ggg aag taa      528
Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys
      165                     170                     175

```

<210> 34

<211> 175

<212> PRT

<213> Erwinia uredovora

<400> 34

Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly
1 5 10 15

Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
20 25 30

Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe
35 40 45

Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu
50 55 60

Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly
65 70 75 80

Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly
85 90 95

Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr
100 105 110

Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
115 120 125

Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser
130 135 140

Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala
145 150 155 160

Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys
 165 170 175

<210> 35

<211> 1520

<212> DNA

<213> Artificial

<220>

<223> Promotor

<400> 35

```

ctcgagtacc gaggcggaac ggcaggaatg ttccctctc ttttagaggg caattcttta      60

tccaatgtca tgttgatgct agatatttct gtctcttata ataaggcgaa taccattttt    120

tgaattgaag ttgagataaa aaaaaagggg gcccaatttg tcaacgcaa agagtcaagc     180

ttttcttttg gctttagccg aacaatctaa gacttattgt ttttgaagat atttgacctt    240

ttctagatat tccttcaagt aaagcttttt tcgagttttt tttttttttc tttgtgaagg    300

atttattggt attggtatcc attttttatt ggaagacaag ataagttaat attgattttg    360

cttaaagatt aaaaggaaat cagaaaacga caataaaaaa tgtaacggac aaactatggt    420

gtcgattata agtctaaatc cttaaaaaat gacaacgagt tgctttcctc tgaaaacaat    480

tcttttgtct ttgcaagaaa ggtttctttt ttgtttgctt gcattactta aacatcaa      540

caaatgaaag gaataaagca gatttgaggg cgaataagga ttttctgggc aacaagatgt    600

gagtgcacc taaggaaact aatgccattc atttgtttta aaacgacatc aaagattgat    660

gatcaacagg attgagagag agaaaaagaa ctcgtgtcat ttatttctgt tgactgaaat    720

tttatattta gaaaaaatgt caaatctata gctttagcta tattacataa catttgaaat    780

aataataata aaaaaagaca cattagagac acttttcaaa ctctaaataa ctgtctataa    840

```

acacaaagaa aacaaagacc tctataacaa cttattagat ttttctcgta cttttgtcta 900
aagatgatgt attcttggtta tcccacactt ctttcatttg ttcttgatgc tactaaatat 960
acaaaatttc ttttttgcaa gagatattat tccaaaaatt ttcaaaaaga aatttttttc 1020
acaatagcag ttgatcgtgt aacccaaaga ggttctttgt tattttgcac ttccgctttg 1080
cggatgatgca tattcaaagt aatatatgga ataaacaacg tgtttaagca tgaaagaaag 1140
gaaacaaagg ccgctttgaa caaatgcata atatttcaga caaaaatgat ctaaagcaag 1200
cagtaaatca aacaagaaac attgctgatt cgcgtagaa aacgataaaa gtctaataag 1260
ccactaagta tacttcaatg aactttttgt atgcttatgg tccaatcaga ccaataattt 1320
gtgaccattc ctgagggtggc tttggtgatg cggaacaga aaaaaatttt ctcaccaatc 1380
gatttaaaaa acaatttctg ctttgaacca aaactttttt tttctcttta atcattaact 1440
ttatcaagta tgtacctacc ctcaaagtc tcaactcaagc acaattatgc taacattggt 1500
ccaccttctc tttagaaatg 1520

<210> 36

<211> 16245

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 36

ccgggctggg	tgccctcgcc	gctgggctgg	cggccgtcta	tggccctgca	aacgcgccag	60
aaacgccgtc	gaagccgtgt	gcgagacacc	gcggccgccg	gcgttgtgga	tacctcgccg	120
aaaacttggc	cctcactgac	agatgagggg	cggacgttga	cacttgaggg	gccgactcac	180
ccggcgcggc	gttgacagat	gaggggcagg	ctcgatttcg	gccggcgacg	tggagctggc	240
cagcctcgca	aatcggcgaa	aacgcctgat	tttacgcgag	tttcccacag	atgatgtgga	300
caagcctggg	gataagtgcc	ctgcggtatt	gacacttgag	gggcgcgact	actgacagat	360
gaggggcgcg	atccttgaca	cttgaggggc	agagtgtctga	cagatgaggg	gcgcacctat	420
tgacatttga	ggggctgtcc	acaggcagaa	aatccagcat	ttgcaagggt	ttccgcccgt	480
ttttcggcca	ccgctaacct	gtcttttaac	ctgcttttaa	accaatattt	ataaaccttg	540
tttttaacca	gggctgcgcc	ctgtgcgcgt	gaccgcgcac	gccgaagggg	ggtgcccccc	600
cttctogaac	cctcccggcc	cgctaacgcg	ggcctcccat	ccccccaggg	gctgcgcccc	660
tcggccgcga	acggcctcac	cccaaaaatg	gcagcgctgg	cagtccttgc	cattgccggg	720
atcggggcag	taacgggatg	ggcgatcagc	ccgagcgcga	cgcccggaag	cattgacgtg	780
ccgcaggtgc	tggcatcgac	attcagcgac	caggtgccgg	gcagtgaggg	cggcggcctg	840
ggtggcggcc	tgcccttcac	ttcggccgtc	ggggcattca	cggacttcat	ggcggggccg	900
gcaattttta	ccttgggcat	tcttggcata	gtggtcgcgg	gtgccgtgct	cgtgttcggg	960
ggtgcgataa	accagcgaa	ccatttgagg	tgataggtaa	gattataccg	aggtatgaaa	1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttataaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaaa cttgcatgga 1260

ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgcgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatat gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgcccggc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcattccaa gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg gggtgagaag cggtgtaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtccggcggg gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gaggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggtttca 2100

aatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggtt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220

aattagcttc ttgggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggttggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc tcaaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc ttatttgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctcgctc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcac caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggcttacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgaag 3360

gagggatgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggg gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaatc atatgggaga agtaccgcaa gctgtcgccg acggcccgc 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatcgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcattct catctcggc gcacttaata tttcgctatt 5100

ctggagcttg ttgtttatth cggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgt ctgctaggta gcccatacgc 5220

attgatggcg gtccctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaagcga gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgccctggcaa ctggcggccg gaggacttct gctcgttoca gtagctttag tgtttgatcc 5460

gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggcttttctca gcccagatc tggggtcgat cagccgggga tgcatacggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttccact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctcgg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgtga cgccgtccc gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcg catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaate 7920

tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaate cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggtg agtaataagc cactggccag 8160

acagctctgg cggtcttgag gtgcagtgga tgattattaa tccgggaccg gccgccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640
tcagtccttg gtaggcagct ttgcccgcgc tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggtcgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttccctt aagtaagtac ttgtctacat ccatactcca tccttcccat cccttatctc 8880
tttgaacctt tcagttcgag ctttcccaact tcctgcgcgc ttgactaaca gctacccgcg 8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgtcttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120
ctacaaagat cgttatgttt atcggcactt tgcctcggcc gcgctccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240
tgtcacgttg caagacctgc ctgaaaccga actgccgct gttctgcgc cggtcgcgga 9300
ggccatggat gcgatcgtg cggccgatct tagccagacg agcgggttcg gcccatctcg 9360
accgaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgct agtgcgtccg tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660
gtggttggtt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcgc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcggttg tgtcgatgtc agtccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaat agtccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttcag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatttg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggc acacttggtt agaggtaatc cttctttcta gctagaagtc ctcggtgtact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt aatctataca 10800

atgctccata gactcacatt gatattgtcg aagatttcga tgctgactta gtagagcaac 10860

tacaaaagtt agcagagaag catgatttct taatctttga agaccgcaag tttgcagata 10920

tcggtatgtg aattctatct attttttttc tgatgtgtgc atggatgact catgatcata 10980

ttcttaggta atactgtcaa gcatcaatat ggcaaggcg tttacaagat tgcttcttgg 11040

tctcatatta ctaatgctca cacagttcct ggagaaggta ttatcaaggg acttgccgaa 11100

gtcggcctcc ctcttggtcg tggcttgctt ttgctagcag aaatgtcatc tcaaggtgca 11160
ttaactaagg gtatttacac tgccgaatct gtcaatatgg ctgccgcaa caaagatttc 11220
gtttttggct ttattgcaca acacaaaatg aatcagtatg atgatgagga ttttggtgtc 11280
atgtcgcttg aagcttggcg taatcatggt catagctgtt tctgtgtga aattgttacc 11340
cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc tgggggtgcct 11400
aatgagttag ctaactcaca ttaattgctg tgcgtcact gccgctttc cagtcgggaa 11460
acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta 11520
ttggggccaaa gacaaaaggc cgacattcaa ccgattgagg gagggaaggc aaatattgac 11580
ggaaattatt cattaaaggc gaattatcac cgtcaccgac ttgagccatt tgggaattag 11640
agccagcaaa atcaccagta gcaccattac cattagcaag gccggaaacg tcaccaatga 11700
aaccatcgat agcagcaccg taatcagtag cgacagaatc aagtttgcct ttagcgtcag 11760
actgtagcgc gttttcatcg gcattttcgg tcatagcccc cttattagcg tttgccatct 11820
tttcataatc aaaatcaccg gaaccagagc caccaccgga accgcctccc tcagagccgc 11880
caccctcaga accgccacc tcagagccac caccctcaga gccgccacca gaaccaccac 11940
cagagccgcc gccagcattg acaggaggcc cgatctagta acatagatga caccgcgcgc 12000
gataatttat cctagtgttc gcgtatat ttgttttcta tcgcgtatta aatgtataat 12060
tgccgggactc taatcataaa aaccatctc ataaataacg tcatgcatta catgttaatt 12120
attacatgct taacgtaatt caacagaaat tatatgataa tcatcgcaag accggcaaca 12180
ggattcaatc ttaagaaact ttattgcaa atgtttgaac gatcggggat catccgggtc 12240
tgtggcgagg actccacgaa aatatccgaa cgcagcaaga tatcgcggtg catctcggtc 12300
ttgcctgggc agtcgccgc gacgccgttg atgtggacgc cgggcccgat catattgtcg 12360

ctcaggatcg tggcgttggtg cttgtcggcc gttgctgtcg taatgatata ggcaccttcg 12420

accgcctggt cgcgagagat cccgtgggcg aagaactcca gcatgagata cccgcgctgg 12480

aggatcatcc agccggcgtc ccggaaaacg attccgaagc ccaacctttc atagaaggcg 12540

gcggtggaat cgaaatctcg tgatggcagg ttgggcgtcg cttggtcggt catttcgaac 12600

cccagagtcc cgctcagaag aactcgtcaa gaaggcgata gaaggcgatg cgctcgcaat 12660

cgggagcggc gataccgtaa agcacgagga agcggtcagc ccattcgccg ccaagctctt 12720

cagcaatata acgggtagcc aacgctatgt cctgatagcg gtccgccaca cccagccggc 12780

cacagtcgat gaatccagaa aagcggccat tttccaccat gatattcggc aagcaggcat 12840

cgccatgggt cagcagcaga tcatcgccgt cgggcattcg cgcccttgagc ctggcgaaca 12900

gttcggctgg cgcgagcccc tgatgtctt cgtccagata atcctgatcg acaagaccgg 12960

cttccatccg agtacgtgct cgctcgatgc gatgtttcgc ttggtggtcg aatgggcagg 13020

tagccggatc aagcgtatgc agccgcgcga ttgcatcage catgatggat actttctcgg 13080

caggagcaag gtgagatgac aggagatcct gccccggcac ttcgcccaat agcagccagt 13140

cccttcccgc ttcagtgaca acgtcgagca cagctgcgca aggaacgccc gtcgtggcca 13200

gccacgatag ccgcgctgcc tcgtcctgca gttcattcag ggcaccggac aggtcggctt 13260

tgacaaaaag aaccgggccc ccctgcgctg acagccggaa cacggcggca tcagagcagc 13320

cgattgtctg ttgtgccag tcatagccga atagcctctc cacccaagcg gccggagaac 13380

ctgcgtgcaa tccatcttgt tcaatcatgc gaaacgatcc agatccggtg cagattatct 13440

ggattgagag tgaatatgag actctaattg gataccgagg ggaatttatg gaacgtcagt 13500

ggagcatttt tgacaagaaa tatttgctag ctgatagtga ccttaggcga cttttgaacg 13560

cgcaataatg gtttctgacg tatgtgctta gtcattaaa ctccagaaac ccgcggtgta 13620

gtggctcctt caacgttgcg gttctgtcag ttccaaacgt aaaacggctt gtcccgcgtc 13680
atcggcgggg gtcataacgt gactccctta attctccgct catgatcaga ttgtcgtttc 13740
ccgccttcag tttaaactat cagtgtttga caggatatat tggcgggtaa acctaagaga 13800
aaagagcggt tattagaata atcggatatt taaaagggcg tgaaaagggt tatccgttcg 13860
tccatttgta tgtgcatgcc aaccacaggg ttcccagat ctggcgccgg ccagcgagac 13920
gagcaagatt ggccgccgcc cgaaacgac cgacagcgcg cccagcacag gtgcgcaggc 13980
aaattgcacc aacgcataca ggcgcagcag aatgccatag tgggcgggtga cgtcgttcga 14040
gtgaaccaga tcgcgcagga ggcccggcag caccggcata atcaggccga tgccgacagc 14100
gtcgagcgcg acagtgtca gaattacgat caggggtatg ttgggtttca cgtctggcct 14160
ccggaccagc ctccgctggt ccgattgaac gcgcggattc tttatcactg ataagttggt 14220
ggacatatta tgtttatcag tgataaagtg tcaagcatga caaagttgca gccgaatata 14280
gtgatccgtg ccgccctgga cctgttgaac gaggtcggcg tagacggtct gacgacagc 14340
aaactggcgg aacggttggg ggttcagcag ccggcgcttt actggcactt caggaacaag 14400
cgggcgctgc tcgacgcact ggccgaagcc atgctggcgg agaatacat gcattcggtg 14460
ccgagagccg acgacgactg gcgctcatth ctgatcgga atgcccgcag cttcaggcag 14520
gcgctgctcg cctaccgcga tggcgcgcg atccatgccg gcacgcgacc gggcgacccg 14580
cagatggaaa cggccgacgc gcagcttcgc ttctctgcg aggcgggttt ttcggccggg 14640
gacgcgctca atgcgctgat gacaatcagc tacttcactg ttggggccgt gcttgaggag 14700
caggccggcg acagcgatgc cggcgagcgc ggcggcaccg ttgaacaggc tccgctctcg 14760
ccgctgttgc gggccgcgat agacgccttc gacgaagccg gtccggacgc agcgttcgag 14820
cagggactcg cggtgattgt cgatggattg gcgaaaagga ggctcgttgt caggaacgtt 14880

gaaggaccga gaaaggggtga cgattgatca ggaccgctgc cggagcgcaa cccactcact 14940
acagcagagc catgtagaca acatccccctc cccctttcca ccgcgtcaga cgcccgtagc 15000
agcccgctac gggctttttc atgccctgcc ctagcgtcca agcctcacgg ccgcgctcgg 15060
cctctctggc ggccttctgg cgctcttccg ctctctcgct cactgactcg ctgcgctcgg 15120
tcgttcgggt gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 15180
aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 15240
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca 15300
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagc actataaaga taccaggcgt 15360
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 15420
tgtccgcctt tctcccttcg ggaagcgtgg cgcttttccg ctgcataacc ctgcttcggg 15480
gtcattatag cgattttttc ggtatatcca tcctttttcg cacgatatac aggattttgc 15540
caaaggggtc gtgtagactt tccttggtgt atccaacggc gtcagccggg caggataggt 15600
gaagtaggcc caccgcgag cgggtgttcc ttcttctactg tcccttattc gcacctggcg 15660
gtgctcaacg ggaatcctgc tctgcgaggc tggccggcta ccgccggcgt aacagatgag 15720
ggcaagcgga tggctgatga aaccaagcca accaggaagg gcagcccacc tatcaagggt 15780
tactgccttc cagacgaacg aagagcgatt gaggaaaagg cggcggcggc cggcatgagc 15840
ctgtcggcct acctgctggc cgtcggccag ggctacaaaa tcacgggcgt cgtggactat 15900
gagcacgtcc gcgagctggc ccgcatcaat ggcgacctgg gccgcctggg cggcctgctg 15960
aaactctggc tcaccgacga cccgcgcaag gcgcggttcg gtgatgccac gatcctcgcc 16020
ctgctggcga agatcgaaga gaagcaggac gagcttggca aggtcatgat gggcgtggtc 16080
cgcccgaggg cagagccatg acttttttag ccgctaaaac ggccgggggg tgcgcgtgat 16140

tgccaagcac gtcccatgc gctccatcaa gaagagcgac ttcgcggagc tgggaagta 16200

catcaccgac gagcaaggca agaccgagcg cctttgcgac gctca 16245

<210> 37

<211> 17877

<212> DNA

<213> Artificial

<220>

<223> Promotor

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 37

ccgggctggt tgccctcgcc gctgggctgg cggcgtcta tggccctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgagg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccggt 480

ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaac cctcccggcc cgctaacgcg ggctcccat cccccaggg gctgcgcccc 660

tcggccgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcatcagc ccgagcgca gcccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840

ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatata gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatactata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgctt aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactcaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgcatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg ttttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcggtccatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtccggcggg gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggtttca 2100

aatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttaagggtt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220

aattagcttc ttggggatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccatth aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttcgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccgcgctg agtcggggca atccccgaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac caccgcgaac aagaaaatcc cgcgcgaggc gctgcaaaac aagggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggg gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggccccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttc 4380

gcctcatgtg cggatcggat tcacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccga cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgccatctt catcctcggc gcacttaata ttctgctatt 5100

ctggagcttg ttgtttattt cggctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtccctgggg ctatttgcg aactgcgggc gtggcgctgt tgggttgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcggga accgtgctga cccgcaagt gcaacctccc gtgcctctgc tcacctttac 5400

cgcttgcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggcttttctca gcccagatc tggggtcgat cagccgggga tgcatacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgga gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cgccgtcccg gactgatggg ctgcctgtat 6060
cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg gggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300
aaatcctggt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggccacta cgtgaacat 6480
cacccaaate aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcattgcac tgaattcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaattcgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttgat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440

tcgaaataca tccgatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcattcgata gctctgcaa gggcgcttgc caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaadc 7920

tgtccagatc atgggtgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggg catgacttga gcctaaaadc cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgtatga aaccggaaag gccgctcagg agctggccag cggcgagac 8580

cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggctcc 8640

tcagtcctg gtaggcagct ttgcccgc tcgtccgccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttgttgc catatcttcc tgctctccc accagctgct 8760

ctttctttt ctctttctt tccatcttc agtatattca tcttccatc caagaacctt 8820

tattcccct aagtaagtac ttgctacat ccatactcca tcttcccat cccttattcc 8880

ttgaacctt tcagttcgag ctttccact tcatcgagc ttgactaaca gctacccgc 8940

ttgagcagac atccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctctgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggg 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacg tctcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acaattgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt ttttcgagtt 10800
tttttttttt ttctttgtga aggatttatt gttattggta tccatttttt attggaagac 10860
aagataagtt aatattgatt ttgcttaaag attaaaagga aatcagaaaa cgacaataaa 10920
aatgtaacg gacaaactat ggtgtcgatt ataagtctaa atccttaaaa aatgacaacg 10980
agttgctttc ctctgaaaac aattcttttg tctttgcaag aaagggttct tttttgtttg 11040
cttgcattac ttaaacatca aatcaaatga aaggaataaa gcagatttga gggcgaataa 11100
ggattttctg gtcaacaaga tgtgagtgac acctaaaggaa ctaaatgcca ttcatttgtt 11160
ttaaacgac atcaaagatt gatgatcaac aggattgaga gagagaaaaa gaactcgtgt 11220
catttatttc tgttgactga aattttatat ttagaaaaaa tgtcaaatct atagctttag 11280
ctatattaca taacatttga aataataata ataaaaaaag acacattaga gacacttttc 11340
aaactctaaa taactgtcta taaacacaaa gaaaacaaag acctctataa caacttatta 11400
gatttttctc gtacttttgt ctaaagatga tgtattcttg ttatcccaca cttctttcat 11460
ttgttcttga tgctactaaa tatacaaaat ttcttttttg caagagatat tattccaaaa 11520
attttcaaaa agaaattttt ttcacaatag cagttgatcg tgtaacccaa agaggttctt 11580
tgttattttg cacttccgct ttgcggtgat gcatattcaa agtaatatat ggaataaaca 11640

acgtgttttaa gcatgaaaga aaggaaacaa aggccgcttt gaacaaatgc ataatatctc 11700

agacaaaaat gatctaaagc aagcagtaaa tcaaacaaga aacattgctg attcgcgcta 11760

gaaaacgata aaagtctaata aagccactaa gtatacttca atgaactttt tgtatgctta 11820

tgggtccaatc agaccaataa tttgtgacca ttcttgaggt ggctttggtg atgcggaaac 11880

agaaaaaaat tttctcacca atcgatttaa aaaacaattt ctgctttgaa ccaaaacttt 11940

ttttttctct ttaatcatta actttatcaa gtatgtacct accctcaaag tcctcactca 12000

agcacaatta tgctaacatt gttccacctt ctctttagaa atgctgtcga agctgcagtc 12060

aatcagcgtc aaggcccgcc gcgttgaact agcccgcgac atcacgcggc ccaaagcttg 12120

cctgcatgct cagcgggtgt cgtagttcg gctgcgagtg gcagcaccac agacagagga 12180

ggcgtgga accgtgcagg ctgccggcg gggcgatgag cacagcgccg atgtagcact 12240

ccagcagctt gaccgggcta tcgcagagcg tcgtgcccg cgcaaacggg agcagctgtc 12300

ataccaggct gccgccattg cagcatcaat tggcgtgtca ggcattgccca tcttcgccac 12360

ctacctgaga tttgccatgc acatgacctt gggcgcgca gtgccatggg gtgaagtggc 12420

tggcactctc ctcttggtgg ttggtggcg gctcggcatg gagatgtatg cccgctatgc 12480

acacaaagcc atctggcatg agtcgcctct gggctggctg ctgcacaaga gccaccacac 12540

acctgcact ggacccttg aagccaacga cttgtttgca atcatcaatg gactgccgc 12600

catgctcctg tgtaccttg gcttctggct gcccaacgtc ctggggcg cctgctttgg 12660

agcggggctg ggcatacgc tatacggcat ggcataatg tttgtacacg atggcctgg 12720

gcacaggcg tttccaccg ggcccatcg tggcctgcc tacatgaagc gcctgacagt 12780

ggcccaccag ctacaccaca gcggcaagta cgggtggcg cctggggta tgttcttgg 12840

tccacaggag ctgcagcaca ttccagggtg ggcggaggag gtggagcgac tggctctgga 12900

actggactgg tccaagcggg agaagcttgg cgtaatcatg gtcatactg tttcctgtgt 12960

gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag 13020

cctgggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgtt 13080

tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 13140

gcggtttgcg tattgggcca aagacaaaag ggcgacattc aaccgattga gggagggaag 13200

gtaaatattg acggaaatta ttcattaaag gtgaattatc accgtcaccg acttgagcca 13260

tttgggaatt agagccagca aaatcaccag tagcaccatt accattagca aggccgaaa 13320

cgtcaccaat gaaaccatcg atagcagcac cgtaatcagt agcgacagaa tcaagtttgc 13380

ctttagcgtc agactgtagc gcgttttcat cggcattttc ggtcatagcc cccttattag 13440

cgtttgccat cttttcataa tcaaaatcac cggaaccaga gccaccaccg gaaccgcctc 13500

cctcagagcc gccaccctca gaaccgccac cctcagagcc accaccctca gagccgccac 13560

cagaaccacc accagagccg ccgccagcat tgacaggagg cccgatctag taacatagat 13620

gacaccgcgc gcgataattt atcctagttt gcgcgctata ttttgttttc tatcgcgat 13680

taaatgtata attgcgggac tctaatacata aaaaccatc tcataaataa cgatcatgcat 13740

tacatgttaa ttattacatg cttaacgtaa ttcaacagaa attatatgat aatcatcgca 13800

agaccggcaa caggattcaa tcttaagaaa ctttattgccc aaatgtttga acgatcgggg 13860

atcatccggg tctgtggcgg gaactccacg aaaatatccg aacgcagcaa gatatcgcg 13920

tgcatctcgg tcttgccctg gcagtcgccg ccgacgccgt tgatgtggac gccggggccc 13980

atcatattgt cgctcaggat cgtggcgctg tgcttgctcg ccgttgctgt cgtaatgata 14040

tcggcacctt cgaccgcctg ttccgcagag atcccgtggg cgaagaactc cagcatgaga 14100

tccccgcgct ggaggatcat ccagccggcg tcccggaaaa cgattccgaa gcccacctt 14160

tcatagaagg cggcgggtgga atcgaaatct cgtgatggca ggttgggcgt cgcttggtcg 14220

gtcatttcga accccagagt cccgctcaga agaactcgtc aagaaggcga tagaaggcga 14280

tgcgctgcga atcgggagcg gcgataccgt aaagcacgag gaagcggcca gccattcgc 14340

cgccaagctc ttcagcaata tcacgggtag ccaacgctat gtcctgatag cggctccgca 14400

caccagccg gccacagtcg atgaatccag aaaagcggcc attttccacc atgatattcg 14460

gcaagcaggc atcgccatgg gtcacgacga gatcatcgcc gtcgggcatg cgcgccctga 14520

gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc ttcgtccaga tcctcctgat 14580

cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat gcgatgttcc gcttggtggg 14640

cgaatgggca ggtagccgga tcaagcgtat gcagccgccc cattgcatca gccatgatgg 14700

atactttctc ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgcccc 14760

atagcagcca gtcccttccc gcttcagtga caacgtcgag cacagctgcg caaggaacgc 14820

ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg cagttcattc agggcaccgg 14880

acaggtcggg cttgacaaaa agaaccgggc gccctgcgc tgacagccgg aacacggcgg 14940

catcagagca gccgattgtc tgttggtgccc agtcatagcc gaatagcctc tccaccaag 15000

cggccggaga acctgcgtgc aatccatctt gttcaatcat gcgaaacgat ccagatccgg 15060

tgcagattat ttggattgag agtgaatatg agactctaata tggataccga ggggaattta 15120

tggaacgtca gtggagcatt tttgacaaga aatatttgct agctgatagt gaccttaggc 15180

gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 15240

accgcgggct gagtggctcc ttcaacgttg cggttctgtc agttccaaac gtaaaacggc 15300

ttgtcccgcg tcatcggcgg gggtcataac gtgactccct taattctccg ctcgatgatca 15360

gattgtcggtt tcccgccttc agtttaaact atcagtgttt gacaggatat attggcgggt 15420

aaacctaaga gaaaagagcg tttattagaa taatcggata tttaaaaggg cgtgaaaagg 15480

tttatccggtt cgtccatttg tatgtgcatg ccaaccacag ggttccccag atctggcgcc 15540

ggccagcgag acgagcaaga ttggccgccc cccgaaacga tccgacagcg cggccagcac 15600

aggtgcgcag gcaaattgca ccaacgcata cagcgccagc agaatgccat agtgggcggt 15660

gacgtcggtt gagtgaacca gatcgcgag gagggccggc agcaccggca taatcaggcc 15720

gatgccgaca gcgtcgagcg cgacagtgtt cagaattacg atcaggggta tgttgggttt 15780

cacgtctggc ctccggacca gcctccgtg gtccgattga acgcgcggat tctttatcac 15840

tgataagttg gtggacatat tatgtttatc agtgataaag tgtcaagcat gacaaagttg 15900

cagccgaata cagtgatccg tgccgccctg gacctgttga acgaggtcgg cgtagacggt 15960

ctgacgacac gcaaactggc ggaacggttg ggggttcagc agccggcgct ttactggcac 16020

ttcaggaaca agcgggcgct gctcgacgca ctggccgaag ccatgctggc ggagaatcat 16080

acgcattcgg tgccgagagc cgacgacgac tggcgctcat ttctgatcgg gaatgcccgc 16140

agcttcaggc aggcgctgct cgcctaccgc gatggcgcgc gcatccatgc cggcacgcga 16200

ccgggcgcac cgcagatgga aacggccgac gcgcagcttc gcttcctctg cgaggcgggt 16260

ttttcgggcg gggacgccgt caatgcgctg atgacaatca gctacttcac tgttggggcc 16320

gtgcttgagg agcaggccgg cgacagcgat gccggcgagc gcggcggcac cgttgaacag 16380

gctccgctct cgccgctggt gcgggcccgc atagacgcct tcgacgaagc cggtcgggac 16440

gcagcgttcg agcagggact cgcggtgatt gtcgatggat tggcgaaaag gaggtcgtt 16500

gtcaggaacg ttgaaggacc gagaaagggt gacgattgat caggaccgct gccggagcgc 16560

aaccactca ctacagcaga gccatgtaga caacatcccc tcccccttc caccgcgtca 16620

gacgcccgta gcagcccgct acgggctttt tcatgccctg ccctagcgtc caagcctcac 16680

ggccgcgctc ggctctctctg gggccttctt ggcgctcttc cgcttcctcg ctcaactgact 16740

cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcaactcaaag gcggtataac 16800

ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 16860

aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg 16920

acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa 16980

gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc 17040

ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgcttttc cgctgcataa 17100

ccctgcttcg gggtcattat agcgattttt tcggtatata catccttttt cgcacgatat 17160

acaggatttt gccaaagggc tcgtgtagac tttccttggc gtatccaacg gcgtcagccg 17220

ggcaggatag gtgaagtagg cccacccgcg agcgggtggt ccttcttcac tgtcccttat 17280

tcgcacctgg cggtgctcaa cgggaatcct gctctgcgag gctggccggc tacgcgggc 17340

gtaacagatg agggcaagcg gatggctgat gaaaccaagc caaccaggaa gggcagccca 17400

cctatcaagg tgtactgcct tccagacgaa cgaagagcga ttgaggaaaa ggcggcgggc 17460

gccggcatga gcctgtcggc ctacctgctg gccgtcggcc agggctacaa aatcacgggc 17520

gtcgtggact atgagcacgt ccgcgagctg gccgcacatca atggcgacct gggccgcctg 17580

ggcggcctgc tgaaactctg gctcacgcac gaccgcgcga cggcgcggtt cggtgatgcc 17640

acgatcctcg ccctgctggc gaagatcgaa gagaagcagg acgagcttgg caaggatcatg 17700

atgggcgtgg tccgcccag ggcagagcca tgactttttt agccgctaaa acggccgggg 17760

ggtgcgcgtg attgccaagc acgtcccat gcgctccatc aagaagagcg acttcgcgga 17820

gctggtgaag tacatcaccg acgagcaagg caagaccgag cgcttttgcg acgctca 17877

<210> 38
<211> 17238
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 38
ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtgga tacctcgccg 120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttcccacag atgatgtgga 300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360
gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccg 480

ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatatth ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaa cctccccggc cgctaacgcg ggctcccat cccccaggg gctgcgcccc 660

tcggccgga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg 720

atcgggggcag taacgggatg ggcgatcagc ccgagcgga cgcccggaag cattgacgtg 780

ccgcagggtgc tggcatcgac attcagcgac cagggtgccg gcagtgagg cggcggcctg 840

ggtggcgcc tgccttcac ttcggccgct ggggcattca cggacttcat ggccgggccc 900

gcaatthta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atattthaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tctthtatat agaagatgc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgctt aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactcaa cttattgata gtgtthtatg ttcagataat gccgatgac tttgtcatgc 1380

agctccaccg atthtgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagctth 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgcatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatat gcgthaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aaatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaagttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220

aattagcttc ttggggatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggtata taagctgggt ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcccttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgctg gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgetgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtccccgttg caggctcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatgtcaa acgctagggc cttgtggggt cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcg cctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgccatctt catcctcggc gcacttaata tttcgctatt 5100

ctggagcttg ttgtttattt cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctdtgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtccctggggg ctatttgctg aactgcgggc gtggcgctgt tgggtgttgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

gcctggcaa ctggcgggcg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttgggtccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcatacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcaact tctaaagaaa tagcgccact cagcttcctc agcgggttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctcg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgca gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cgccgtcccg gactgatggg ctgcctgtat 6060
cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300
aatcctgtt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggg 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc caggggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcgggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttggg gcatagaggg 7320

tcctttggct ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaata 7920

tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggg catgacttga gcctaaaatc cgccgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100
tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgagac 8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640
tcagtccttg gtaggcagct ttgccccgct tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggctggtgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880
tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctacccgc 8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcggtgcttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120
ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcttgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcggttg tgtcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaat agtccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctctgttact 10740
gtgtaagcgc ccaactccaca tctccactcg acctgcaggc atgcaagctt ctaccgcttg 10800
gaccagtcca gttccaggac cagtcgctcc acctcctccg ccgcacctgg aatgtgctgc 10860
agctcctgtg gacccaagaa cataccccag ggcgogccac cgtacttgcc gctgtggtgt 10920
agctggtggg ccaactgtcag gcgcttcag tagggcaggc cagcgatggg cccggtggga 10980
aagcgctgtg gcaccaggcc atcgtgtaca aacatatatg ccatgccgta tagcgtgatg 11040
cccagccccg ctccaaagca ggccgcccc aggacgttg gcagccagaa gccaaaggta 11100
cacaggagca tggcgggcag tccattgatg attgcaaaca agtcgttggc ttcaaaggtt 11160
ccagtgcgag gtgtgtggtg gctcttgtgc agcagccagc ccagaggcga ctcatgccag 11220
atggctttgt gtgcatagcg ggcatacatc tccatgccga gcgcgccacc aaccaccaag 11280
aggagagtgc cagccacttc accccatggc actgcgccgc ccacgggtcat gtgcatggca 11340
aatctcaggt aggtggcgaa gatggcaatg cctgacacgc caattgatgc tgcaatggcg 11400
gcagcctggt atgacagctg ctcccgtttg cgccgggcac gacgctctgc gatagcccg 11460
tcaagctgct ggagtgtac atcggcgctg tgctcatcgc ccgcgcggc agcctgcacg 11520
gttcccagcg cctcctctgt ctgtggtgct gccactcgca gccgaactaa cgagcaccgc 11580
tgagcatgca ggcagacttt gggccgcgtg atgtcgcggg ctagttcaac gcggcgggcc 11640
ttgacgctga ttgactgcag cttcgacagc atagagataa aataaaaaga gaagaaaaga 11700
aagtttgtac aatttctttt tgtttatata acatacacgc tatgtcaaca tttagaataa 11760

gggggaaaaa atcttccatc atattogaat gcacaagatt atttctttgt tcgctctttt 11820
tggtcgggtc atcgagattt agagtgtaat caaagatact gtcattctga gagcgttgca 11880
caggctgctg ttgccaaat tggatgtttg ccgaattagt aaaatacgca agcatttctt 11940
acctttccgc tcccttttcc taattctccc aaagactaaa tgaggaaaga taaaggacaa 12000
agaaaatgta aagacaaaga aattgaaaac gatataaact tgcagcacgt aagaccaaag 12060
caaattggtg actattcttg tgtacaaaca tgtataaaaa aaaacttttt tttgctcctg 12120
gaggacaaaa tttcaaactc cttgaagaag attgcttgta tatctatcat atgcatatat 12180
catatcgatg gaaaaagaaa gtcaggcatg tatttataaa aagaagaatg tgccatgctt 12240
ccgaatttct tttcactttc ttttccttat ctattttaat ctcaagcttg gcgtaatcat 12300
ggatcatagct gtttcctgtg tgaaattggt atccgctcac aattccacac aacatacgag 12360
ccggaagcat aaagtgtaaa gcctgggggtg cctaattgagt gagctaactc acattaattg 12420
cgttgcgctc actgccgctt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 12480
tcggccaacg cgcggggaga ggcggtttgc gtattgggcc aaagacaaaa gggcgacatt 12540
caaccgattg agggagggaa ggtaaattatt gacggaaatt attcattaaa ggtgaattat 12600
caccgtcacc gacttgagcc atttggaat tagagccagc aaaatcacca gtagcaccat 12660
taccattagc aaggccggaa acgtcaccaa tgaaaccatc gatagcagca ccgtaatcag 12720
tagcgacaga atcaagtttg ccttttagcgt cagactgtag cgcgttttca tcggcatttt 12780
cggatcatagc ccccttatta gcgtttgcca tcttttcata atcaaaatca ccggaaccag 12840
agccaccacc ggaaccgcct ccctcagagc cgccaccctc agaaccgcca ccctcagagc 12900
caccaccctc agagccgcca ccagaaccac caccagagcc gccgccagca ttgacaggag 12960
gccgatcta gtaacataga tgacaccgag cgcgataatt tatcctagtt tgcgcgctat 13020

at tt t t g t t t t c t a t c g c g t a t t a a t g t a t a a t t g c g g g a c t c t a a t c a t a a a a c c c a t 13080

c t c a t a a a t a a c g t c a t g c a t t a c a t g t t a a t t a t t a c a t g c t t a a c g t a a t t c a a c a g a 13140

a a t t a t a t g a t a a t c a t c g c a a g a c c g g c a a c a g g a t t c a a t c t t a a g a a a c t t t a t t g c 13200

c a a a t g t t t g a a c g a t c g g g g a t c a t c c g g g t c t g t g g g c g g g a a c t c c a c g a a a t a t c c 13260

g a a c g c a g c a a g a t a t c g c g t g c a t c t c g t c t t g c c t g g g c a g t c g c c g c c g a c g c c g 13320

t t g a t g t g g a c g c c g g g c c c g a t c a t a t t g t c g t c a g g a t c g t g g g c g t t g t g c t t g t c g 13380

g c c g t t g c t g t c g t a a t g a t a t c g g c a c c t t c g a c c g c c t g t t c c g c a g a g a t c c c g t g g 13440

g c g a a g a a c t c c a g c a t g a g a t c c c c g c g c t g g a g g a t c a t c c a g c c g g c g t c c c g g a a a 13500

a c g a t t c c g a a g c c c a a c c t t t c a t a g a a g g c g g g c g g t g g a a t c g a a a t c t c g t g a t g g c 13560

a g g t t g g g c g t c g c t t g g t c g g t c a t t t c g a a c c c a g a g t c c c g c t c a g a a g a a c t c g t 13620

c a a g a a g g c g a t a g a a g g c g a t g c g c t g c g a a t c g g g a g c g g c g a t a c c g t a a a g c a c g a 13680

g g a a g c g g t c a g c c c a t t c g c c g c c a a g c t c t t c a g c a a t a t c a c g g g t a g c c a a c g c t a 13740

t g t c c t g a t a g c g g t c c g c c a c a c c a g c c g g c c a c a g t c g a t g a a t c c a g a a a g c g g c 13800

c a t t t t c c a c c a t g a t a t t c g g c a a g c a g g c a t c g c c a t g g g t c a c g a c g a g a t c a t c g c 13860

c g t c g g g c a t g c g c g c c t t g a g c c t g g c g a a c a g t t c g g c t g g c g c g a g c c c c t g a t g c t 13920

c t t c g t c c a g a t c a t c c t g a t c g a c a a g a c c g g c t t c c a t c c g a g t a c g t g t c g c t c g a 13980

t g c g a t g t t t c g c t t g g t g g t c g a a t g g g c a g g t a g c c g g a t c a a g c g t a t g c a g c c g c c 14040

g c a t t g c a t c a g c c a t g a t g a t a c t t t t c t c g g c a g g a g c a a g g t g a g a t g a c a g g a g a t 14100

c c t g c c c c g g c a c t t c g c c c a a t a g c a g c c a g t c c c t t c c c g c t t c a g t g a c a a c g t c g a 14160

g c a c a g c t g c g c a a g g a a c g c c c g t c g t g g c c a g c c a c g a t a g c c g c g c t g c c t c g t c c t 14220

g c a g t t c a t t c a g g g c a c c g g a c a g g t c g g t c t t g a c a a a a a g a a c c g g g c g c c c c t g c g 14280

ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttggtgcc cagtcatagc 14340
cgaatagcct ctccacccaa gcggccggag aacctgcgtg caatccatct tgttcaatca 14400
tgcgaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14460
ttggataccg aggggaatth atggaacgtc agtggagcat ttttgacaag aaatatttgc 14520
tagctgatag tgaccttagg cgacttttga acgcgcaata atggtttctg acgtatgtgc 14580
ttagctcatt aaactccaga aaccgcggc tgagtggctc cttcaacgtt gcggttctgt 14640
cagttccaaa cgtaaaacgg cttgtccgc gtcacggcg ggggtcataa cgtgactccc 14700
ttaattctcc gctcatgac agattgtcgt ttccgcctt cagtttaaac tatcagtgtt 14760
tgacaggata tattggcggg taaacctaag agaaaagagc gtttattaga ataatcggat 14820
atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gccaacca 14880
gggttcccca gatctggcg cggccagcga gacgagcaag attggccgcc gcccgaaacg 14940
atccgacagc gcgccagca cagggtgcga ggcaaattgc accaacgcat acagcgccag 15000
cagaatgcca tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg 15060
cagcaccggc ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac 15120
gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15180
aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa 15240
gtgtcaagca tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg 15300
aacgaggtcg gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt gggggttcag 15360
cagccggcgc ttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa 15420
gccatgctgg cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca 15480
tttctgatcg ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg 15540

cgcatccatg ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt 15600
cgcttcctct gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc 15660
agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag 15720
cgcgggcgga ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc 15780
ttcgacgaag ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga 15840
ttggcgaaaa ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga 15900
tcaggaccgc tgccggagcg caaccactc actacagcag agccatgtag acaacatccc 15960
ctcccccttt ccacgcgctc agacgcccg agcagccgc tacgggcttt ttcatgccct 16020
gccctagcgt ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt 16080
ccgcttcctc gctcactgac tcgtgcgct cggtcgttcg gctgcggcga gcggtatcag 16140
ctcactcaaa ggcgtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 16200
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcttt 16260
tccataggct ccgccccct gacgagcatc aaaaaatcg acgctcaagt cagagggtggc 16320
gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc ctctgctcgt 16380
ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctcct tcgggaagcg 16440
tggcgctttt ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat 16500
ccatcctttt tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg 16560
tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt 16620
tcctttctca ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga 16680
ggctggccgg ctaccgcgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag 16740
ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16800

attgaggaaa aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc 16860
cagggctaca aaatcacggg cgctgtggac tatgagcacg tccgcgagct ggcccgcac 16920
aatggcgacc tggggccgct gggcggcctg ctgaaactct ggctcacga cgaccgcgc 16980
acggcgcggt tcggtgatgc cacgacctc gccctgctgg cgaagatcga agagaagcag 17040
gacgagcttg gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt 17100
tagccgctaa aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat 17160
caagaagagc gacttcgcg agctggtgaa gtacatcacc gacgagcaag gcaagaccga 17220
gcgcctttgc gacgctca 17238

<210> 39

<211> 17238

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 39

ccgggctggt tgcctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccc gcgttggtga tacctcgcg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgagg gccgactcac 180

ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcgcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgagg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaa cctcccggcc cgctaacgcg ggctcccat cccccagg gctgcgcccc 660

tcggccgcga acggcctcac ccaaaaatg gcagcgtgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgca cggccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac cagggtgccg gcagtgagg cgcgggcctg 840

ggtggcgccc tgccttcac ttcgccgctc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatactata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttccggag actgtcatat gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggg gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggtttca 2100
aaatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaagggtt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220
aattagcttc ttggggatat tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaaggata taagctgggt ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc ttaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc ttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gggggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgaag 3360
gagggtgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatggtoc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660
tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720
agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
ccctgttcac caccgcgaac aagaaaatcc cgcgcgaggc gctgcaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggg gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgctgtg ctgtttgctg 4260
gcgaccacta caccgaaatc atatgggaga agtaccgcaa gctgtcgccg acggcccgcac 4320
ggatgttoga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttc 4380
gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcattgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg caccgcgcgc ttacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggccttg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860
catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata ttctcgtatt 5100

ctggagcttg ttgtttattt cgggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccatacag 5220

attgatggcg gtctctgggg ctatttgcgg aactgcgggc gtggcgctgt tgggtttgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgcttgga ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc ttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcacaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcaact tctaaagaaa tagcgccact cagcttctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgtga cgcgctccc gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggatccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcca 6900

gtcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtgc gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaatata tccggatgtc gaaggcttgg ggcacctgcy ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaate 7920

tgccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga ccagagggt catgacttga gcctaaaate cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgagg agagacggac ggacgcagag agaagggtg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640

tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggctgttgc gtcagtccaa ctttgttgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac ttgtctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcggtcttt 9060

cagcttcgat gtagggggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgca ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcacaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgct agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcggtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctggggtt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttccctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt agagataaaa 10800
taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac atacacgcta 10860
tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc acaagattat 10920
ttctttgttc gctctttttg gtccgggtcat cgagatttag agtgtaatca aagatactgt 10980
catctcgaga gcgttgacaca ggctgctggt tgccaaattg gatgtttgcc gaattagtaa 11040
aatacgcaag catttcttac ctttccgctc ccttttccta attctccaa agactaaatg 11100
aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga tataaacttg 11160
cagcacgtaa gaccaaagca aattggtaac tattcttgtg tacaacatg tataaaaaaa 11220
aacttttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat tgcttgtata 11280
tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta ttataaaaa 11340

gaagaatgtg ccatgcttcc gaatttcttt tcactttctt ttccttatct attttaatct 11400
catgctgtcg aagctgcagt caatcagcgt caaggcccg cgcgttgaac tagcccgca 11460
catcacgcgg cccaaagtct gcctgcatgc tcagcgggtgc tcgttagttc ggctgcgagt 11520
ggcagcacca cagacagagg aggcgctggg aaccgtgcag gctgccggcg cgggcgatga 11580
gcacagcgcc gatgtagcac tccagcagct tgaccgggct atcgagagc gtcgtgccc 11640
gcgcaaacgg gagcagctgt cataccaggc tgccgccatt gcagcatcaa ttggcgtgtc 11700
aggcattgcc atcttcgcca cctacctgag atttgccatg cacatgaccg tgggcggcgc 11760
agtgccatgg ggtgaagtgg ctggcactct cctcttggtg gttggtggcg cgctcggcat 11820
ggagatgtat gcccgtatg cacacaaagc catctggcat gagtgcctc tgggctggct 11880
gctgcacaag agccaccaca cacctgcac tggaccctt gaagccaacg acttgtttgc 11940
aatcatcaat ggactgccc ccatgctcct gtgtacctt ggcttctggc tgcccaacgt 12000
cctgggggcg gcctgctttg gagcggggct gggcatcacg ctatacggca tggcatatat 12060
gtttgtacac gatggcctgg tgcacaggcg ctttcccacc gggcccatcg ctggcctgcc 12120
ctacatgaag cgcctgacag tggccacca gctacaccac agcggcaagt acggtggcgc 12180
gccctggggt atgttcttgg gtccacagga gctgcagcac attccaggtg cggcggagga 12240
ggtggagcga ctggtcctgg aactggactg gtccaagcgg tagaagcttg gcgtaatcat 12300
ggtcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag 12360
ccggaagcat aaagtgtaaa gcctgggggtg cctaattgagt gagctaactc acattaattg 12420
cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 12480
tcggccaacg cgcggggaga ggcggtttgc gtattgggccc aaagacaaaa gggcgacatt 12540
caaccgattg agggagggaa ggtaaatatt gacggaaatt attcattaaa ggtgaattat 12600

caccgtcacc gacttgagcc atttggaat tagagccagc aaaatcacca gtagcaccat 12660
taccattagc aaggccggaa acgtcaccaa tgaaaccatc gatagcagca ccgtaatcag 12720
tagcgacaga atcaagtttg ccttttagcgt cagactgtag cgcgttttca tcggcatttt 12780
cggtcatagc ccccttatta gcgtttgccca tcttttcata atcaaaatca ccggaaccag 12840
agccaccacc ggaaccgcct ccctcagagc cgccaccctc agaaccgccca ccctcagagc 12900
caccaccctc agagccgccca ccagaaccac caccagagcc gccgccagca ttgacaggag 12960
gcccgatcta gtaacataga tgacaccgcg cgcgataatt taccctagtt tgcgcgctat 13020
attttgtttt ctatcgcgta ttaaattgtat aattgcggga ctctaatacat aaaaacccat 13080
ctcataaata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13140
aattatatga taatcatcgc aagaccggca acaggattca atcttaagaa actttattgc 13200
caaatgtttg aacgatcggg gatcatccgg gtctgtggcg ggaactccac gaaaatatcc 13260
gaacgcagca agatatcgcg gtgcattctg gtcttgcccg gccagtcgcc gccgacgccg 13320
ttgatgtgga cgccggggccc gatcatattg tcgctcagga tcgtggcggt gtgcttgctg 13380
gccgttgctg tcgtaatgat atcggcacct tcgaccgcct gttccgcaga gatcccggtg 13440
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccggaag 13500
acgattccga agcccaacct ttcatagaag gcggcggtgg aatcgaaatc tcgtgatggc 13560
aggttgggcg tcgcttggtc ggtcatttcg aaccccagag tcccgtcag aagaactcgt 13620
caagaaggcg atagaaggcg atgcgctgcg aatcgggagc ggcgataccg taaagcacga 13680
ggaagcggtc agccattcgc ccgccaagct cttcagcaat atcacgggta gccaacgcta 13740
tgtcctgata gcggtcgcc acaccagcc ggccacagtc gatgaatcca gaaaagcggc 13800
cattttccac catgatattc ggcaagcagg catcgccatg ggtcacgacg agatcatcgc 13860

cgtcggggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc ccctgatgct 13920
cttcgtccag atcatcctga tcgacaagac cggttccat ccgagtacgt gctcgtcga 13980
tgcatgttt cgcttggtgg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc 14040
gcattgcac agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat 14100
cctgccccgg cacttcgccc aatagcagcc agtcccttcc cgcttcagtg acaacgtcga 14160
gcacagctgc gcaaggaacg cccgtcgtgg ccagccacga tagccgcgt gcctcgtcct 14220
gcagttcatt cagggcaccg gacaggtcgg tcttgacaaa aagaaccggg cgccctgcg 14280
ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttggtgcc cagtcatagc 14340
cgaatagcct ctccacccaa gcggccggag aacctgcgtg caatccatct tgttcaatca 14400
tgcgaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14460
ttggataccg aggggaattt atggaacgtc agtggagcat ttttgacaag aaatatttgc 14520
tagctgatag tgaccttagg cgacttttga acgcgaata atggtttctg acgtatgtgc 14580
ttagctcatt aaactccaga aaccgcggc tgagtggctc cttcaacgtt gcggttctgt 14640
cagttccaaa cgtaaaacgg cttgtcccgc gtcacggcg ggggtcataa cgtgactccc 14700
ttaattctcc gctcatgatc agattgtcgt tccccgctt cagtttaaac tatcagtgtt 14760
tgacaggata tattggcggg taaacctaa agaaaagagc gtttattaga ataatcgat 14820
atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gccaccaca 14880
gggttcccc a gatctggcg cgccagcga gacgagcaag attggccgcc gcccgaaacg 14940
atccgacagc gcgccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag 15000
cagaatgcc tagtggggcg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg 15060
cagcaccggc ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac 15120

gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15180

aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa 15240

gtgtcaagca tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg 15300

aacgaggtcg gcgtagacgg tctgacgaca cgcaaactgg cggaacgggt ggggggttcag 15360

cagccggcgc ttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa 15420

gccatgctgg cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca 15480

tttctgatcg ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg 15540

cgcatccatg ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgcagctt 15600

cgcttctctc gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc 15660

agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag 15720

cgcgggcgca ccgttgaaca ggctccgctc tcgccgctgt tgccggccgc gatagacgcc 15780

ttcgacgaag ccggtccgga cgcagcgctc gagcaggac tcgcggtgat tgtcgatgga 15840

ttggcgaaaa ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga 15900

tcaggaccgc tgccggagcg caacccactc actacagcag agccatgtag acaacatccc 15960

ctcccccttt ccaccgcgtc agacgcccg agcagccgc tacgggcttt ttcatgccct 16020

gccctagcgt ccaagcctca cggccgcgct cggcctctct ggccggcctc tggcgctctt 16080

ccgcttcctc gctcactgac tcgctgcgct cggctgctcg gctgcggcga gcggtatcag 16140

ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 16200

tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcttt 16260

tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagagggtggc 16320

gaaaccgcac aggactataa agataccagg cgtttcccc tggaagctcc ctcgtgcgct 16380

ctcctgttcc gacctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg 16440
tggcgctttt ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat 16500
ccatcctttt tcgcacgata tacaggattt tgccaaaggg ttctgttaga ctttccttgg 16560
tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt 16620
tccttcttca ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga 16680
ggctggccgg ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag 16740
ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16800
attgaggaaa aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc 16860
cagggttaca aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc 16920
aatggcgacc tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgaccgcgc 16980
acggcgcggt tcggtgatgc cagcatcctc gccctgctgg cgaagatcga agagaagcag 17040
gacgagcttg gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt 17100
tagccgctaa aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat 17160
caagaagagc gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga 17220
gcgcctttgc gacgtca 17238

<210> 40

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature
<222> (3471)..(3471)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3679)..(3679)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3770)..(3770)
<223> n is a, c, g, or t

<400> 40
gatctttcga cactgaaata cgtcgagcct gtcgcgcttg gaagcggcga ggagcctcgt 60
cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct 240
tcgggaagcc agcgactggg gacctccagc gtccctttaa ggctgccaac agctttctca 300
gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360
gtggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg 420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480
tttgatttg ctttccaggc tgagactcta gcttgagca tagagggtcc tttggcttctc 540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga 600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720
tccgatagct ctgcaaaggc cgttgcacaa tgcaagtcaa acgttgctag cagttccagg 780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840

caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg 900

cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960

tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca 1020

gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140

gttgaccggt gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200

ggagtgacct agagggtcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260

aaatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcacgggaga atatggagct 1500

tcacgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560

cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgac tcgtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgg taagctccct 1680

aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccgg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccttggtg 1860

ggcagctttg ccccgctctgt ccgccgggtg tgcggcggg gttgacaagg tcgttgcgtc 1920

agtccaacat ttgttgccat attttctcgc tctccccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040

taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100

gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160
accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
gtctccgacc tgatgcagct ctccgagggc gaagaatctc gtgctttcag cttcgatgta 2280
ggagggcgctg gatatgtcct gggggtaaat agctgcgccg atggtttcta caaagatcgt 2340
tatgtttatc ggcactttgc atcgccgcgcg ctcccgattc cggaagtgtc tgacattggg 2400
gaattcagcg agagcctgac ctattgcatc tcccgccgtg cacaggggtg cacgttgcaa 2460
gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg 2520
atcgtgcgg ccatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580
ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640
tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700
atgctttggg ccgaggactg ccccgaaagtc cggcacctcg tgcacgcgga tttcggctcc 2760
aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820
ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgagg atcgccgcgg 2940
ctccgggcgt atatgctccg cattggctct gaccaactct atcagagctt ggttgacggc 3000
aatctcgatg atgcagcttg ggcgcagggc cgatgcgacg caatcgtccg atccggagcc 3060
gggactgtcg ggcgtacaca aatcgccgcg agaagcgcg cgtctggac cgatggctgt 3120
gtagaagtac tcgccgatag tggaaaccga cggccagca ctggtccgag ggcaaaggaa 3180
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240
tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300
tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ccttctagt 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480
gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540
tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600
atatattgtct ctaattgtac ttgacatgc tctcttctt tactctgata gcttgactat 3660
gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720
ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttcctactaa 3780
gatggtatac aatagtaacc atgcatggtt gcctagttaa tgctccgtaa cacccaatac 3840
gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggtaca 3900
cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960
ctccacatct ccactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020
agaaagttaa aagaaattcg gaagcatggc acattcttct ttttataaat acatgcctga 4080
ctttcttttt ccatcgatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140
ggagtttgaa attttgtcct ccaggagcaa aaaaaagttt ttttttatac atgtttgtac 4200
acaagaatag ttaccaattt gctttggtct tacgtgctgc aagtttatat cgttttcaat 4260
ttctttgtct ttacattttc ttgttccttt atctttcctc atttagtctt tgggagaatt 4320
aggaaaaggg agcggaaagg taagaaatgc ttgcgtatth tactaattcg gcaaacatcc 4380
aatttggcaa acagcagcct gtgcaacgct ctcgagatga cagtatctt gattacactc 4440
taaattctga tgacccgacc aaaaagagcg aacaaagaaa taatcttgtg cattcgaata 4500
tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560
aaaaagaaat tgtacaaact ttcttttctt ctctttttat tttatctcta tgctgtogaa 4620

gctgcagtca atcagcgtca aggcccgccg cgttgaacta gcccgcgaca tcacgcgggc 4680

caaagtctgc ctgcatgctc agcgggtgctc gttagttcgg ctgcgagtgg cagcaccaca 4740

gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg gccgatgagc acagcgccga 4800

tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccggc gcaaacggga 4860

gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920

cttcgccacc tacctgagat ttgccatgca catgaccgtg ggccggcgag tgccatgggg 4980

tgaagtggct ggcactctcc tcttgggtgt tggtggcgcg ctggcatgg agatgtatgc 5040

ccgctatgca cacaagcca tctggcatga gtgcctctg ggctggctgc tgcacaagag 5100

ccaccacaca cctcgactg gaccctttga agccaacgac ttgtttgcaa tcatcaatgg 5160

actgccccgc atgctcctgt gtacctttgg cttctggctg cccaacgtcc tggggggggc 5220

ctgctttgga gcgggggtgg gcatcacgct atacggcatg gcatatatgt ttgtacacga 5280

tggcctgggt cacaggcgt tcccaccgg gcccatcgct ggctgccct acatgaagcg 5340

cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgcg cctgggggtat 5400

gttcttgggt ccacaggagc tgcagcacat tccagggtcg gcggaggagg tggagcgact 5460

ggtcctggaa ctggactggt ccaagcggta gattgtgact gatagcgaga ctctgggtcg 5520

atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaaatac agcaaggcaa 5580

cggccgtagc ctaggtgatc aaagactgtt gggcttgtct ctgaagcttg taggaaaggc 5640

agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700

cgggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760

gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaaga ttgctctgtt 5820

gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880

gattaagagt ttttttgttg attagtgaat ttttgtttta tttatgtctg atagttcaat 5940

aaagagacaa cacatacaat ataaaatcat tgtcttttaa tgtaattta gtagagtgt 6000

aagcctgcat ttttttgtga cgcataaaca atgaattcac cccgcttctg gtttttaa 6060

aattatgtca aactagggaa aattcttttt tttctcttcg ttcttttttt ggcttgttgt 6120

ggagtacag gcttgtcttc agattgatag aggttgtata cactcaacag agcaatcttg 6180

gcacgttcgc ttcttttttag atgagctctt gtaggatatg ttgtcttctg ctgcttgagt 6240

acagatccaa ttgcagcata tacttgga gacgctcgta caccgccttg acagtttgcc 6300

ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tctacaagc 6360

ttcagagaca agccaacag tctttgatca cctaggctac gggcggtgcc ttgctgtatt 6420

tgttctgttt ctctttttct aagccattgt tgaggcagat aacatcgacc caacatcctc 6480

gagccatact acagcataaa aggatacgtt ttctttaaca gaaatttacc cttttgttat 6540

cagcacatac aaaaaaaaaa aaatttaaga tgagtaggac ttccattctc tcaaaaattt 6600

tattcaatcc ataatgaat tatttttggg caaaaaagaa agattatgcc tgattttctc 6660

tatttttttt ttttttacia ctcaccaat actttctagc ccagcttgcc gtaatcatgg 6720

tcatagctgt ttctgtgtg aaattgttat ccgctcacia ttccacacia catagcagcc 6780

ggaagcataa agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaattgcg 6840

ttgcgctcac tgcccgttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 6900

ggccaacgcg cggggagagg cggtttgcgt attgggcaa agacaaaagg gcgacattca 6960

accgattgag ggaggggaagg taaatattga cggaaattat tcattaaagg tgaattatca 7020

ccgtcaccga cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080

ccattagcaa ggccggaaac gtcaccaatg aaaccatoga tagcagcacc gtaatcagta 7140

gcgacagaat caagtttgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg 7200
gtcatagccc ccttatttagc gtttgccatc ttttcataat caaaatcacc ggaaccagag 7260
ccaccaccgg aaccgcctcc ctgagagccg ccaccctcag aaccgccacc ctgagagcca 7320
ccaccctcag agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggc 7380
ccgatctagt aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat 7440
tttgttttct atcgcgtatt aaatgtataa ttgcgggact ctaatcataa aaacccatct 7500
cataaataac gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560
ttatatgata atcatcgcaa gaccggcaac aggattcaat cttaagaaac tttattgcca 7620
aatgtttgaa cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga 7680
acgcagcaag atatcgcggt gcatctcggt ctgacctggg cagtcgccgc cgacgccgtt 7740
gatgtggacg cggggcccga tcatattgtc gctcaggatc gtggcggtgt gcttgctggc 7800
cgttgctgtc gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc 7860
gaagaactcc agcatgagat ccccgcgctg gaggatcatc cagccggcgt cccggaaaac 7920
gattccgaag cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag 7980
gttgggcgtc gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca 8040
agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100
aagcggtcag cccattcgcc gccaaagtct tcagcaatat cacgggtagc caacgctatg 8160
tcctgatagc ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcgcca 8220
ttttccacca tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg 8280
tcgggcatgc gcgccttgag cctggcgaac agtcgggctg gcgcgagccc ctgatgctct 8340
tcgtccagat catcctgatc gacaagaccg gcttccatcc gactacgtgc tcgctcgatg 8400

cgatgttttcg cttggtggtc gaatgggcag gtagccggat caagcgtatg cagccgccgc 8460

attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520

tgccccggca cttcgcccaa tagcagccag tcccttcccg cttcagtgc aacgtcgagc 8580

acagctgcgc aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc 8640

agttcattca gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct 8700

gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgcca gtcatagccg 8760

aatagcctct ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820

cgaaacgac cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880

ggataccgag ggggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940

gctgatagt accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000

agctcattaa actccagaaa cccgcggtc agtggtcct tcaacgttgc ggttctgtca 9060

gttccaaacg taaaacggct tgtcccgct catcgggggg ggtcataacg tgactccctt 9120

aattctccgc tcatgatcag attgtcggtt cccgccttca gtttaaacta tcagtgtttg 9180

acaggatata ttggcgggta aacctagag aaaagagcgt ttattagaat aatcgatat 9240

ttaaaagggc gtgaaaagg ttatccgttc gtccatttgt atgtgcatgc caaccacagg 9300

gttccccaga tctggcgccg gccagcgaga cgagcaagat tggccgccgc ccgaaacgat 9360

ccgacagcgc gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca 9420

gaatgccata gtggcggtg acgtcgttcg agtgaaccag atcgcgcagg agggccggca 9480

gcaccggcat aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga 9540

tcaggggtat gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa 9600

cgcgcggatt ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt 9660

gtcaagcatg acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa 9720
cgaggtcggc gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca 9780
gccggcgctt tactggcact tcaggaacaa gggggcgctg ctcgacgcac tggccgaagc 9840
catgctggcg gagaatcata cgcattcggt gccgagagcc gacgacgact ggcgctcatt 9900
tctgatcggg aatgcccgca gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg 9960
catccatgcc ggcacgcgac cggggcgacc gcagatggaa acggccgacg cgcagcttcg 10020
cttcctctgc gaggcgggtt tttcggccgg ggacgccgtc aatgcgctga tgacaatcag 10080
ctacttcact gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg 10140
cggcggcacc gttgaacagg ctccgctctc gccgctgttg cgggccgca tagacgcctt 10200
cgacgaagcc ggtccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt 10260
ggcgaaaagg aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc 10320
aggaccgctg ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct 10380
ccccctttcc accgcgtcag acgcccgtag cagcccgtc cgggcttttt catgccctgc 10440
cctagcgctc aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc 10500
gcttcctcgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 10560
cactcaaagg cggtaatcag gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 10680
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 10740
aaccgcagag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 10800
cctgttccga ccctgccgct taccggatac ctgtccgctt ttctcccttc gggaagcggtg 10860
gcgcttttcc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc 10920

atcctttttc gcacgatata caggattttg ccaaagggtt cgtgtagact ttccttggtg 10980
tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccacccgcga gcgggtgttc 11040
cttcttcact gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg 11100
ctggccggct accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160
aaccaggaag ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat 11220
tgaggaaaag gcggcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 11280
gggctacaaa atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 11340
tggcgacctg ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac 11400
ggcgcgggtc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 11460
cgagcttggc aaggtcatga tgggcgtggt ccgcccaggg gcagagccat gactttttta 11520
gccgctaaaa cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca 11580
agaagagcga cttcgcggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc 11640
gcctttgcga cgctcaccgg gctggttgcc ctcgccgctg ggctggcggc cgtctatggc 11700
cctgcaaacg cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt 11760
tgtggatacc tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact 11820
tgaggggccc actcaccggc gcggcgcttg acagatgagg ggcaggctcg atttcggccc 11880
gcgacgtgga gctggccagc ctcgcaaac gcgaaaaacg cctgatttta cgcgagtttc 11940
ccacagatga tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc 12000
gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 12060
tgaggggcgc acctattgac atttgagggg ctgtccacag gcagaaaac cagcatttgc 12120
aagggtttcc gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca 12180

atatttataa accttgtttt taaccagggc tgcgcctgt gcgctgacc gcgcacgccg 12240
aaggggggtg ccccccttc tcgaaccctc ccggcccgt aacgcgggcc tcccatcccc 12300
ccagggggtg cggccctcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt 12360
ccttgccatt gccgggatcg gggcagtaac gggatgggcg atcagcccga gcgcgacgcc 12420
cggaagcatt gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag 12480
tgagggcggc ggcctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga 12540
cttcatggcg gggccggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc 12600
cgtgctcgtg ttccgggggtg cgataaacc agcgaaccat ttgaggtgat aggtaagatt 12660
ataccgaggt atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat 12720
ttaaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg cttgtaat 12780
attgacaata ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga 12840
tttcaggggg caaggcatag gcagcgcgt tatcaatata tctatagaat gggcaaagca 12900
taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt 12960
ctatcataat tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc 13020
gatgactttg tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg 13080
tgccaggtgc tgccctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct 13140
gattacgtgc agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca 13200
tatcaccacg tcaaaggggtg acagcaggct cataagacgc ccagcgtcg ccatagtgcg 13260
ttcaccgaat acgtgcgcaa caaccgtctt ccggagactg tcatacgcg aaaacagcca 13320
gcgctggcgc gatttagccc cgacatagcc ccactgttcg tccatttccg cgagacgat 13380
gacgtcactg cccggctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga 13440

cgtaaaatcg tgttgaggcc aacgcccata atgcgggctg ttgcccgga tccaacgcca 13500
ttcatggcca tatcaatgat tttctggtgc gtaccgggtt gagaagcggg gtaagtgaac 13560
tgcagttgcc atgttttacg gcagtgagag cagagatagc gctgatgtcc ggcggtgctt 13620
ttgccgttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680
gccactggag cacctcaaaa acaccatcat acactaaatc agtaagttgg cagcatcacc 13740
cataattgtg gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaanaac 13800
aactttgaaa aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg 13860
gagttcgtct tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa 13920
ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980
accgctgcgt aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag 14040
aaaatgaaaa cctatattta aaaatgacgg acagccggta taaagggacc acctatgatg 14100
tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaaggctc 14160
tgcactttga acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc 14220
tttgctcgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280
agtgcacag gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag 14340
acagccgctt agcgaattg gattacttac tgaataacga tctggccgat gtggattgag 14400
aaaactggga agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga 14460
cggaaaagcc cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct 14520
ttgtgaaaga tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca 14580
agtggatatga cattgccttc tgcgtccggt cgatcaggga ggatatacggg gaagaacagt 14640
atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt 14700

atattttact ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag 14760
caggagcgca ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc 14820
aagtatttgg gcaaggggtc gctgggtattc gtgcagggca agattcggaa taccaagtac 14880
gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940
gacaccaagg caccaggcgg gtcaaatacag gaataagggc acattgcccc ggcgtgagtc 15000
ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc ata'caggcaa 15060
gaactgatcg acgcgggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc 15120
atgctgtcgc cccgcgaaac cttccagtcg gtccggctcga tggtcagca agctacggcc 15180
aagatcgagc gcgacagcgt gcaactggct cccctgccc tgcccgcgcc atcgcccgcc 15240
gtggagcgtt cgcgtcgtct cgaacaggag gcggcaggtt tggcgaagtc gatgaccatc 15300
gacacgcgag gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa 15360
caggtcagcg aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa 15420
atgcagcttt ccttgttcga tattgcgccg tggccggaca cgatgcgagc gatgccaac 15480
gacacggccc gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg 15540
caaaacaagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag 15600
ctgcggggccg acgatgacga actggtgtgg cagcaggtgt tggagtacgc gaagcgcacc 15660
cctatcggcg agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggctg 15720
atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780
atgggcttca cgtccgaccg cgttgggcac ctggaatcgg tgtcgctgct gcaccgcttc 15840
cgcgtcctgg accgtggcaa ga'aaacgtcc cgttgccagg tcctgatcga cgaggaaatc 15900
gtcgtgctgt ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg 15960

tcgccgacgg cccgacggat gttcgactat ttcagctcgc accgggagcc gtaccgcctc 16020

aagctggaaa ccttcgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc 16080

gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg 16140

gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg 16200

ggttcagcag ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact 16260

tgcttcgctc agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag 16320

gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 16380

aggatttccg cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg 16440

tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500

tggcattcgg cgctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg 16560

acggcccca gacgctcac aaggcgcac tgtccggcgt tttcgtggag cccgaacagc 16620

gaggccgagg ggtcgccggt atgctgctgc gggcggtgcc ggcgggttta ttgctcgtga 16680

tgatcgctccg acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac 16740

ttaatatctc gctattctgg agcttggtgt ttatttcggt ctaccgcctg ccgggcgggg 16800

tcgcggcgac ggtaggcgt gtgcagcgc tgatggctgt gttcatctct gccgctctgc 16860

taggtagccc gatacgattg atggcggtcc tgggggctat ttgcggaact gcgggcgtgg 16920

cgctgttggt gttgacacca aacgcagcgc tagatcctgt cggcgctcga gcgggcctgg 16980

cgggggcggt ttccatggcg ttcggaaccg tgctgacctg caagtggcaa cctcccgtgc 17040

ctctgctcac ctttaccgcc tggcaactgg cgccggagg acttctgctc gttccagtag 17100

ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggcctggcgt 17160

ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220

tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280
cggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcgggtgag 17340
caatggatag gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc 17400
ttctcagcg gctttatcca gcgatttcct attatgtcgg catagttctc aagatcgaca 17460
gcctgtcacg gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg 17520
agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct 17580
ccgcgagatc atcogtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc 17640
ggtaacatga gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccgact 17700
gatgggctgc ctgtatcgag tggtgatttt gtgccgagct gccggtcggg gagctgttgg 17760
ctggctgggtg gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac 17820
acattgcgga cgtttttaat gtactgggggt ggtttttctt ttcaccagtg agacgggcaa 17880
cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggt 17940
ttgccccagc aggcgaaaat cctgtttgat ggtggttccg aaatcgcaa aatcccttat 18000
aaatcaaaag aatagcccga gatagggttg agtggtgttc cagtttggaa caagagtcca 18060
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120
ccactacgtg aaccatcacc caaatcaagt tttttggggg cgaggtgccg taaagcacta 18180
aatcggaacc ctaaaggag ccccgattt agagcttgac ggggaaagcc ggcgaaactg 18240
gcgagaaaagg aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg 18300
ggaaggcgga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagttggg taacgccagg gttttccag tcacgacgtt gtaaaacgac 18420
ggccagtga ttcgagctcg gtacccggg 18449

<210> 41
<211> 18449
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (3471)..(3471)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3679)..(3679)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3770)..(3770)
<223> n is a, c, g, or t

<400> 41
gatcttttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60

cctgtcacaa ctaccaacat ggagtaacat aagggccagt tccgccagct cattaagagc 120

cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180

cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct 240

tcgggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agctttctca 300

gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360

gtggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg 420

atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480

tttggatttg ctttccaggc tgagactcta gcttgagca tagagggtcc tttggctttc 540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga 600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720
tccgatagct ctgcaaaggc cgttgcacaa tgcaagtcaa acgttgctag cagttccagg 780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840
caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg 900
cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960
tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca 1020
gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080
ctatacagaa taagataggt ggagagctta taccgagctc ccaaatctgt ccagatcatg 1140
gttgaccggt gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200
ggagtgacct agagggtcat gacttgagcc taaaatccgc cgcctccacc attttagaa 1260
aaatgtgacg aactcgtgag ctctgtacag tgaccgggtga ctctttctgg catgcggaga 1320
gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380
ctctgagggt cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440
aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcacggaga atatggagct 1500
tcacgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560
cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgatc tggtaaaaga 1620
ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680
aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctgt ccgcccgggtg tgcggcgagg gttgacaagg tcgttgcgtc 1920

agtccaacat ttgttgccat attttctctg tctccccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040

taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100

gttcgagctt tcccacttca tcgcagcttg actaacagct acccgcttg agcagacatc 2160

accatgcctg aactcaccgc gacgtctgtc gagaagtctc tgatcgaaaa gttcgacagc 2220

gtctccgacc tgatgcagct ctcgaggggc gaagaatctc gtgctttcag cttcgatgta 2280

ggagggcggtg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcgggccgc ctcccgattc cggaagtgtc tgacattggg 2400

gaattcagcg agagcctgac ctattgcac tcccgccgtg cacagggtgt cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgaggaggc catggatgcg 2520

atcgtgcgg ccgatcttag ccagacgagc gggttcggcc cattcgagc gcaaggaatc 2580

ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640

tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700

atgctttggg ccgaggactg ccccgaaagtc cggcacctcg tgcacgcgga ttccggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcggta ttgactggag cgaggcgatg 2820

ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880

atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgcagg atcgccgcgg 2940

ctccgggcgt atatgctccg cattggctctt gaccaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggc cgatgcgacg caatcgccg atccggagcc 3060
gggactgtcg ggcgtacaca aatcgcccgc agaagcgcgg ccgtctggac cgatggctgt 3120
gtagaagtac tcgccgatag tggaaccga cgcgccagca ctcgccgag ggcaaaggaa 3180
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240
tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300
tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagt ctttctagt 3360
atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480
gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540
tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggctcc acgcgactat 3600
atatttgtct ctaattgtac tttgacatgc tcctcttctt tactctgata gcttgactat 3660
gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720
ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttcctactaa 3780
gatggtatac aatagtaacc atgcatggtt gcctagtga tgctccgtaa cacccaatac 3840
gccggccgaa acttttttac aactctcta tgagtcgttt acccagaatg cacaggtaca 3900
cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgcca 3960
ctccacatct cactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020
agaaagtga aagaaattcg gaagcatggc acattcttct tttataaat acatgcctga 4080
ctttcttttt ccatcgatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140
ggagtttgaa attttgcct ccaggagcaa aaaaaagttt tttttatac atgtttgtac 4200
acaagaatag ttaccaattt gctttggtct tacgtgctgc aagtttatat cgttttcaat 4260

ttctttgtct ttacattttc ttgttccttt atctttcctc atttagtctt tgggagaatt 4320

aggaaaaggg agcggaagg taagaaatgc ttgcgtatct tactaattcg gcaaaccatcc 4380

aatttggcaa acagcagcct gtgcaacgct ctcgagatga cagtatcttt gattacactc 4440

taaatctcga tgacccgacc aaaaagagcg aacaaagaaa taatcttgtg cattogaata 4500

tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560

aaaaagaaat tgtacaaact ttcttttctt ctctttttat tttatctcta tgctgtcgaa 4620

gctgcagtca atcagcgtca aggcccgcg cgttgaacta gcccgcgaca tcacgcggcc 4680

caaagtctgc ctgcatgctc agcgggtgctc gttagtctcg ctgcgagtgg cagcaccaca 4740

gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg ggcgatgagc acagcgccga 4800

tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccgcc gcaaaccggga 4860

gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920

cttcgccacc tacctgagat ttgccatgca catgaccgtg ggcggcgagc tgccatgggg 4980

tgaagtggct ggcactctcc tcttgggtgt tgggtggcgcg ctccggcatgg agatgtatgc 5040

ccgctatgca caaaaagcca tctggcatga gtgcctctg ggctggctgc tgcacaagag 5100

ccaccacaca cctcgactg gaccctttga agccaacgac ttgtttgcaa tcatcaatgg 5160

actgcccgcc atgctcctgt gtacctttgg cttctggctg cccaacgtcc tgggggcggc 5220

ctgctttgga gcggggctgg gcatcacgct atacggcatg gcatatatgt ttgtacacga 5280

tggcctgggt cacaggcgct tcccaccgg gcccatcgct ggctgcctt acatgaagcg 5340

cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgcg cctgggggtat 5400

gttcttgggt ccacaggagc tgcagcacat tccagggtcg gcggaggagg tggagcgact 5460

ggtcctggaa ctggactggt ccaagcgggc gattgtgact gatagcgaga ctctgggtcg 5520

atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaaatac agcaaggcaa 5580

cgcccgtagc ctaggtgatac aaagactggt gggcttgtct ctgaagcttg taggaaaggc 5640

agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700

cgggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760

gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaaga ttgctctggt 5820

gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880

gattaagagt ttttttgttg attagtgaat ttttgtttta tttatgtctg atagttcaat 5940

aaagagacaa cacatacaat ataaaatcat tgtctttaaa tgtaattta gtagagtgt 6000

aagcctgcat ttttttgtg cgcataaaca atgaattcac cccgcttctg gtttttaaat 6060

aattatgtca aactagggaa aattcttttt tttctcttcg ttcttttttt ggcttggtgt 6120

ggagtcacag gcttgcttc agattgatag aggttgata cactcaacag agcaatcttg 6180

gcacgttcgc ttcttttag atgagctctt gtaggatatg ttgtcttctg ctgcttgagt 6240

acagatccaa ttgcagcata tacttgga gacgctcgta caccgccttg acagtttgcc 6300

ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tcctacaagc 6360

ttcagagaca agcccaacag tctttgatca cctaggctac gggcggtgccc ttgctgtatt 6420

tggtctgttt cttcttttct aagccattgt tgaggcagat aacatcgacc caacatcctc 6480

gagccatact acagcataaa aggatacggt ttctttaaca gaaatttacc cttttgttat 6540

cagcacatac aaaaaaaaaag aaatttaaga tgagtaggac ttccattctc tcaaaaaattt 6600

tattcaatcc ataatgaat ttttttggg caaaaaagaa agattatgcc tgattttctc 6660

tatttttttt ttttttacia ctccaccaat actttctagc ccagcttggc gtaatcatgg 6720

tcatagctgt ttctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc 6780

ggaagcataa agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaatgtcg 6840

ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 6900

ggccaacgcg cggggagagg cggtttgcgt attgggcaa agacaaaagg gcgacattca 6960

accgattgag ggagggagagg taaatattga cggaaattat tcattaaagg tgaattatca 7020

ccgtcaccga cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080

ccattagcaa ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta 7140

gcgacagaat caagtttgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg 7200

gtcatagccc ccttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag 7260

ccaccaccgg aaccgcctcc ctgagagccg ccaccctcag aaccgccacc ctgagagcca 7320

ccaccctcag agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggc 7380

ccgatctagt aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat 7440

tttgttttct atcgcggtatt aaatgtataa ttgcgggact ctaatcataa aaacccatct 7500

cataaataac gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560

ttatatgata atcatcgcaa gaccggcaac aggattcaat cttagaaac tttattgcca 7620

aatgtttgaa cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga 7680

acgcagcaag atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt 7740

gatgtggacg ccgggcccga tcatattgtc gtcaggatc gtggcgttgt gcttgtcggc 7800

cgttgctgtc gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc 7860

gaagaactcc agcatgagat ccccgcgctg gaggatcatc cagccggcgt cccggaaaac 7920

gattccgaag cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag 7980

gttgggcgtc gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca 8040

agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100
aagcggtcag ccatttcgcc gccaaagtct tcagcaatat cacgggtagc caacgctatg 8160
tcctgatagc ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca 8220
ttttccacca tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg 8280
tcgggcatgc gcgccttgag cctggcgaa agttcggctg gcgcgagccc ctgatgctct 8340
tcgtccagat catcctgac gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg 8400
cgatgtttcg cttgggtggc gaatgggcag gtagccggat caagcgtatg cagccgccgc 8460
attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520
tgccccggca cttcgcccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc 8580
acagctgcgc aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc 8640
agttcattca gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg ccctgcgct 8700
gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgcca gtcatagccg 8760
aatagcctct ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820
cgaaacgac cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880
ggataccgag gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940
gctgatagtg accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000
agctcattaa actccagaaa ccgcggcgtg agtggctcct tcaacgttgc ggttctgtca 9060
gttccaaaacg taaaacggct tgtccgcgt catcgccggg ggtcataacg tgaactccctt 9120
aattctccgc tcatgatcag attgtcgttt ccgccttca gtttaaacta tcagtgtttg 9180
acaggatata ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcggatat 9240
ttaaaagggc gtgaaaaggc ttatccgttc gtccatttgt atgtgcatgc caaccacagg 9300

gttccccaga tctggcgccg gccagcgaga cgagcaagat tggccgccc cggaaacgat 9360

ccgacagcgc gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca 9420

gaatgccata gtggcggtg acgtcgttcg agtgaaccag atcgcgcagg aggccccgca 9480

gcaccggcat aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga 9540

tcaggggtat gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa 9600

cgcgcggtt ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt 9660

gtcaagcatg acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa 9720

cgaggtcggc gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca 9780

gccggcgctt tactggcact tcaggaacaa gggggcgctg ctgacgcac tggccgaagc 9840

catgctggcg gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt 9900

tctgatcggg aatgcccgca gcttcaggca ggcgctgtc gcctaccgag atggcgcgcg 9960

catccatgcc ggcacgcgac cgggcgaccc gcagatggaa acggccgacg cgcagcttcg 10020

cttcctctgc gaggcggtt tttcgccggg ggacgccgtc aatgcgctga tgacaatcag 10080

ctacttcact gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg 10140

cggcggcacc gttgaacagg ctccgtctc gccgctgtg cgggcccga tagacgcctt 10200

cgacgaagcc ggtccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt 10260

ggcgaaaagg aggtcgttg tcaggaacgt tgaaggaccg agaaagggtg acgattgatc 10320

aggaccgctg ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct 10380

ccccctttcc accggtcag acgcccgtag cagcccgtc cgggcttttt catgccctgc 10440

cctagcgtcc aagcctcac gccgctcg gcctctctgg cggccttctg gcgctcttcc 10500

gcttctctgc tactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 10560

cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620

tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 10680

cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 10740

aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 10800

cctgttccga ccctgccgt taccggatac ctgtccgcct ttctcccttc ggaagcgtg 10860

gcgctttttc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc 10920

atcctttttc gcacgatata caggattttg ccaaaggggt cgtgtagact ttccttggtg 10980

tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccaccgcga gcgggtgttc 11040

cttcttcact gtcccttatt cgcacctggc ggtgctcaac gggaatctg ctctgcgagg 11100

ctggccggct accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160

aaccaggaag ggcagccac ctatcaagg gtactgcctt ccagacgaac gaagagcgat 11220

tgaggaaaag gggcgggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 11280

gggctacaaa atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 11340

tggcgacctg ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac 11400

ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 11460

cgagcttggc aaggatcatga tgggcgtggt ccgcccagg gcagagccat gactttttta 11520

gccgctaaaa cgccgggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca 11580

agaagagcga cttcgcgagg ctggtgaagt acatcaccga cgagcaaggc aagaccgagc 11640

gcctttgcga cgctaccgg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc 11700

cctgcaaacg cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgcccggct 11760

tgtggatacc tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact 11820

tgaggggccc actcaccgcc cgcggcgttg acagatgagg ggcaggctcg atttcggccc 11880

gcgacgtgga gctggccagc ctcgcaaadc ggcgaaaacg cctgatttta cgcgagtttc 11940

ccacagatga tgtggacaag cctggggata agtgccttgc ggtattgaca cttgaggggc 12000

gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 12060

tgaggggcgc acctattgac atttgagggg ctgtccacag gcagaaaadc cagcattttgc 12120

aagggtttcc gcccgttttt cgccaccgc taacctgtct tttaacctgc ttttaaacca 12180

atatttataa accttgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacgccg 12240

aaggggggtg ccccccttc tcgaaccctc ccggcccgct aacgcggggc tcccatcccc 12300

ccaggggctg cgccctcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt 12360

ccttgccatt gccgggatcg gggcagtaac gggatgggag atcagcccga gcgcgacgcc 12420

cggaagcatt gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag 12480

tgagggcggc ggctgggtg gcggcctgcc cttacttcg gccgtcgggg cattcacgga 12540

cttcattggc ggccgggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc 12600

cggtgctctg ttcgggggtg cgataaaccc agcgaaccat ttgaggtgat aggtgaagatt 12660

ataccgaggt atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat 12720

ttaaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg ctttgaatat 12780

attgacaata ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga 12840

tttcaggggg caaggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca 12900

taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt 12960

ctatcataat tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgccc 13020

gatgactttg tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccc 13080

tgccaggtgc tgcctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct 13140
gattacgtgc agctttocct tcaggcggga ttcatacagc ggccagccat ccgtcatcca 13200
tataccacg tcaaaggggtg acagcaggct cataagacgc cccagcgtcg ccatagtgcg 13260
ttcaccgaat acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca 13320
gcgctggcgc gatttagccc cgacatagcc ccaactgttcg tccatttccg cgcagacgat 13380
gacgtcactg cccggctgta tgcgcgaggt taccgactgc ggctgagtt ttttaagtga 13440
cgtaaaatcg tgttgaggcc aacgcccata atgcgggctg ttgcccggca tccaacgcca 13500
ttcatggcca tatcaatgat tttctggtgc gtaccgggtt gagaagcggg gtaagtgaac 13560
tgcagttgcc atgttttacg gcagtgcgag cagagatagc gctgatgtcc ggcggtgctt 13620
ttgccgttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680
gccactggag cacctcaaaa acaccatcat acactaaatc agtaagttgg cagcatcacc 13740
cataattgtg gtttcaaaat cggctccgtc gatactatgt tatacgcaa ctttgaaaac 13800
aactttgaaa aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg 13860
gagttcgtct tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa 13920
ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980
accgctgcgt aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag 14040
aaaatgaaaa cctatatatta aaaatgacgg acagccggta taaagggacc acctatgatg 14100
tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtcc 14160
tgcactttga acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgctc 14220
tttgctcgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280
agtgcacgag gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag 14340

acagccgctt agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg 14400

aaaactggga agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaga 14460

cggaaaagcc cgaagaggaa cttgtctttt occacggcga cctgggagac agcaacatct 14520

ttgtgaaaga tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca 14580

agtggtatga cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt 14640

atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt 14700

atattttact ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag 14760

caggagcgca ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc 14820

aagtatttgg gcaaggggtc gctgggtattc gtgcagggca agattcggaa taccaagtac 14880

gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940

gacaccaagg caccaggcgg gtcaaatacag gaataagggc acattgcccc ggcgtagtc 15000

ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060

gaactgatcg acgcggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc 15120

atgcgtgctc cccgcgaaac cttccagtc gtcggctcga tgggtccagca agctacggcc 15180

aagatcgagc gcgacagcgt gcaactggct cccctgccc tgcccgcgcc atcggccgcc 15240

gtggagcggt cgcgtcgtct cgaacaggag gcggcagggt tggcgaagtc gatgaccatc 15300

gacacgcgag gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa 15360

caggtcagcg aggccaaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa 15420

atgcagcttt ccttgttcga tattgcgccg tggccggaca cgatgcgagc gatgccaaac 15480

gacacggccc gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg 15540

caaaacaagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgctcgag 15600

ctgcgggccc acgatgacga actggtgtgg cagcaggtgt tggagtacgc gaagcgcacc 15660

cctatcggcg agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggtcg 15720

atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780

atgggcttca cgtccgaccg cgttgggcac ctggaatcgg tgcgctgct gcaccgcttc 15840

cgcgtcctgg accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc 15900

gtcgtgctgt ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg 15960

tcgccgacgg cccgacggat gttcgactat ttcagctcgc accgggagcc gtaccgctc 16020

aagctggaaa ccttccgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc 16080

gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg 16140

gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg 16200

ggttcagcag ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact 16260

tgcttcgctc agtatcgctc gggacgcacg gcgcgtcta cgaactgccg ataaacagag 16320

gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 16380

aggatttccg cgagatccga ttgtcgccc tgaagaaagc tccagagatg ttcgggtccg 16440

tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500

tggcattcgg cgcctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg 16560

acggccccaa ggacgctcac aaggcgcac tcgtccggcgt tttcgtggag cccgaacagc 16620

gaggccgagg ggtcgccggt atgctgctgc ggcggttgcc ggcggttcta ttgctcgtga 16680

tgatcgccg acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac 16740

ttaatatctt gctattcttg agcttggtgt ttatttcggt ctaccgcctg ccgggcgggg 16800

tcggggcgac ggtaggcgct gtgcagccgc tgatggctgt gttcatctct gccgctctgc 16860

taggtagccc gatacgattg atggcgggtcc tgggggctat ttgcggaact gcgggcgtgg 16920

cgctgttggt gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg 16980

cgggggcggt ttccatggcg ttcggaaccg tgctgaccog caagtggcaa cctcccgtgc 17040

ctctgctcac ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag 17100

ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggctggcgt 17160

ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220

tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280

cggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcgggtgag 17340

caatggatag gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc 17400

ttcctcagcg gctttatcca gcgatttcct attatgtcgg catagtcttc aagatcgaca 17460

gcctgtcacg gttaagcgag aaatgaataa gaaggctgat aattcggatc tctcggaggg 17520

agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct 17580

ccgcgagatc atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc 17640

ggtaacatga gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccgact 17700

gatgggctgc ctgtatcgag tgggtgatttt gtgccgagct gccggtcggg gagctgttgg 17760

ctggctgggtg gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac 17820

acattgcgga cgtttttaat gtactggggg ggtttttctt ttcaccagtg agacgggcaa 17880

cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggg 17940

ttgccccagc aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat 18000

aaatcaaaag aatagcccga gatagggttg agtgtgttgc cagtttgga caagagtcca 18060

ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120

ccactacgtg aaccatcacc caaatcaagt tttttgggggt cgaggtgccg taaagcacta 18180
aatcggaacc ctaaaggag cccccgattt agagcttgac ggggaaagcc ggogaacgtg 18240
gcgagaaagg aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg 18300
ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagtggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 18420
ggccagtga ttcgagctcg gtacccggg 18449

<210> 42
<211> 17593
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 42
ccgggctggg tgcctcgcc gctgggctgg cggcgtcta tggcctgca aacgcgccag 60
aaacgcgctc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgcg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcgggcga aacgcctgat ttacgcgag tttccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcgcca ccgtaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgctg gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaa cctcccgccc cgctaacgcg ggctcccat cccccaggg gctgcgcccc 660

tcggccgga acggcctcac ccaaaaaatg gcagcgtgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgga cgcccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840

ggtggcgccc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcga ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatc gcggtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac caggtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgcccggc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcattcaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aaatcggtc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggtt tagaatgcaa ggaacagtga attggagtgc gtcttgttat 2220

aattagcttc ttggggatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaagggtata taagctgggt ggagaaaatg aaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc ttattgatc ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940

cttctgcgtc cggctgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggctcgtggc attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcggtgcaaa tcaggaataa gggcacattg ccccgcgctg agtcggggca atcccgaag 3360

gaggggtgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggg gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa togggtgtcg tgetgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgc 4320

ggatgttcga ctatttcagc togcaccggg agccgtaccg gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcattgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860

catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgctatt 5100

ctggagcttg ttgtttatct cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtcctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcggg accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgcttgga ctggcgccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcatacaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctcgc agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940

gtttcaaacc cggcagctta gttgcggtt tccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgtga cgccgtccc gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cggccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020

aagatccaag tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtggg ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaatt 7920
tgtccagatc atgggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980
cctagctgat tctggagtga cccagagggg catgacttga gcctaaaatc cgccgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100
tggcatgcgg agagacggac ggacgcagag agaagggtg agtaataagc cactggccag 8160
acagctctgg cggtcttgag gtgcagtgga tgattattaa tccgggaccg gccgccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggtg cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640
tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggtcgttgc gtcagtccaa catttgttgc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttatcc 8880
tttgaacctt tcagttcgag ctttcccact tcatcgagc ttgactaaca gctacccgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctctgtcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgtcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgccgct gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgtgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtcaa gcagcaaaga 10140

gtgccttota gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncagggt cggcggaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctggggt cgcaaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgcgggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctctgttact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt ttttcgagtt 10800

tttttttttt ttctttgtga aggatttatt gttattggta tccatttttt attggaagac 10860

aagataagtt aatattgatt ttgcttaaag attaaaagga aatcagaaaa cgacaataaa 10920

aaatgtaacg gacaaactat ggtgtcgatt ataagtctaa atccttaaaa aatgacaacg 10980

agttgctttc ctctgaaaac aattcttttg tctttgcaag aaaggtttct tttttgtttg 11040

cttgcatcac ttaaacaatca aatcaaatga aaggaataaa gcagatttga gggcgaataa 11100

ggattttctg gtcaacaaga tgtgagtgac acctaggaa ctaaagcca ttcatttggt 11160

ttaaaacgac atcaaagatt gatgatcaac aggattgaga gagagaaaaa gaactcgtgt 11220

catttatttc tgttgactga aattttatat ttagaaaaaa tgtcaaatct atagctttag 11280

ctatattaca taacatttga aataataata ataaaaaag acacattaga gacacttttc 11340

aaactctaaa taactgtcta taaacacaaa gaaaacaaag acctctataa caacttatta 11400

gattttttctc gtacttttgt ctaaagatga tgtattcttg ttatcccaca cttctttcat 11460

ttgttcttga tgctactaaa tatacaaaat ttcttttttg caagagatat tattccaaaa 11520
atthttcaaaa agaaatthttt ttcacaatag cagttgatcg tgtaacccaa agaggthctt 11580
tgthattttg cacttccgct ttgcggtgat gcatattcaa agtaatatat ggaataaaca 11640
acgtgtthta gcatgaaaga aaggaaacaa aggccgcttt gaacaaatgc ataathttt 11700
agacaaaaat gatctaaagc aagcagtaaa tcaaacaaga aacattgctg attcgcgthta 11760
gaaaacgata aaagtctaata agccactaa gtatacttca atgaacttht tgtatgctta 11820
tggtccaatc agaccaataa thtgtgacca thcttgaggt ggctthtggtg atgcggaaac 11880
agaaaaaaat thttctacca atcgattthta aaaacaattt ctgctthgaa ccaaaactth 11940
ththththctt thaatcatta actthtatcaa gtatgtacct accctcaaag thctcactca 12000
agcacaatta tgctaacatt gthccacctt ctctthtagaa atgtthtgga thtggaatgc 12060
cttgatcgth thcgthaccg tgattggcat ggaagtgatt gctgcactgg cacacaaata 12120
catcatgcac ggctggggth ggggatggca thththcat catgaaccgc gtaaaggthc 12180
gthtgaaagth aacgatctth atgccgtggt ththtgctga thatcgatcc tgctgattth 12240
thtgggcagth acaggaatgt ggccgctcca gtggattggc gcaggthtga cggcgthtgg 12300
athactctat ththtggtgc acgacgggct gthgcatcaa cgtthggcat thcgctatat 12360
thcacgcaag ggctacctca aacggtthta tatggcgcac cgtatgcath acgccgthcag 12420
gggcaaagaa gthtgththt ththtggtt thcttatgag cggccctgt caaaactthca 12480
ggcgacgctc cgggaaagac atggcgctag agcgggcgct gccagagatg cgcagggcg 12540
ggaggatgag cccgcatccg ggaagtaagg gcctgaccag aggcggccag cagcagcgth 12600
aathththcg gctggtgctg tgactgccgc tgatcccaaa gctthggcgt atcatgthca 12660
tagctgtht ctgthtgaaa thgttatccg thcacaattc cacacaacat acgagccgga 12720

agcataaagt gtaaagcctg gggcgcctaa tgagtgcgct aactcacatt aattgcgttg 12780
cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc 12840
caacgcgcgg ggagaggcgg tttgcgtatt gggccaaaga caaaaggcgg acattcaacc 12900
gattgagggg ggaaggtaa atattgacgg aaattattca ttaaagggtga attatcacgg 12960
tcaccgactt gagccatttg ggaattagag ccagcaaaat caccagtagc accattacca 13020
ttagcaaggc cggaaacgtc accaatgaaa ccatcgatag cagcacgta atcagtagcg 13080
acagaatcaa gtttgccttt agcgtcagac tgtagcgcgt tttcatcggc attttcggtc 13140
atagccccct tattagcgtt tgccatcttt tcataatcaa aatcacggga accagagcca 13200
ccaccggaac cgcctccctc agagccgcca ccctcagaac cgccaccctc agagccacca 13260
ccctcagagc cgccaccaga accaccacca gagccgcccgc cagcattgac aggaggcccg 13320
atctagtaac atagatgaca ccgcgcgcga taatttatcc tagtttgccg gctatatattt 13380
gttttctatc gcgtattaaa tgtataattg cgggactcta atcataaaaa cccatctcat 13440
aaataacgtc atgcattaca tgttaattat tacatgctta acgtaattca acagaaatta 13500
tatgataatc atcgcaagac cggcaacagg attcaatctt aagaaacttt attgccaaat 13560
gtttgaacga tcggggatca tccgggtctg tggcgggaac tccacgaaaa tatccgaacg 13620
cagcaagata tcgcggtgca tctcggtctt gcctgggcag tcgccgccga cgccgttgat 13680
gtggacgccg ggcccgatca tattgtcgct caggatcgtg gcgttggtgt tgtcggccgt 13740
tgctgtcgta atgatatcgg caccttcgac cgctgttcc gcagagatcc cgtgggcgaa 13800
gaactccagc atgagatccc cgcgctggag gatcatccag ccggcgtccc ggaaaacgat 13860
tccgaagccc aacctttcat agaaggcggc ggtggaatcg aaatctcgtg atggcagggt 13920
gggcgtcgct tggtcggtca tttcgaacct cagagtcccc ctcagaagaa ctcgtcaaga 13980

aggcgataga aggcgatgcg ctgcgaatcg ggagcggcga taccgtaaag cacgaggaag 14040
cggtcagccc attcgccgcc aagctcttca gcaatatcac gggtagccaa cgctatgtcc 14100
tgatagcggc cgcgcacacc cagccggcca cagtcgatga atccagaaaa gcggccattt 14160
tccaccatga tattcggcaa gcaggcatcg ccatgggtca cgacgagatc atcgccgtcg 14220
ggcatgcgcg ccttgagcct ggcgaaacagt tcggctggcg cgagcccctg atgctcttcg 14280
tccagatcat cctgatcgac aagaccggct tccatccgag tacgtgctcg ctcgatgcga 14340
tgtttcgctt ggtggctgaa tgggcaggta gccggatcaa gcgtatgcag ccgccgcatt 14400
gcatcagcca tgatggatac tttctcggca ggagcaaggc gagatgacag gagatcctgc 14460
cccggcactt cgcccaatag cagccagtcc cttcccgctt cagtgacaac gtcgagcaca 14520
gctgcgcaag gaacgcccgt cgtggccagc cacgatagcc gcgctgcctc gtcttcgagt 14580
tcattcaggg caccggacag gtcggctcttg acaaaaagaa ccgggcgccc ctgcgctgac 14640
agccggaaca cggcggcatc agagcagccg attgtctgtt gtgcccagtc atagccgaat 14700
agcctctcca cccaagcggc cggagaacct gcgtgcaatc catcttggtc aatcatgcga 14760
aacgatccag atccggtgca gattatcttg attgagagtg aatatgagac tctaattgga 14820
taccgagggg aatttatgga acgtcagtg agcatctttg acaagaaata tttgctagct 14880
gatagtgacc ttaggcgact tttgaacgcg caataatggt ttctgacgta tgtgcttagc 14940
tcattaaact ccagaaaccc gcggctgagt ggctccttca acgttgcggt tctgtcagtt 15000
ccaaacgtaa aacggcttgt cccgcgtcat cggcgggggt cataacgtga ctcccttaat 15060
tctccgctca tgatcagatt gtcgtttccc gccttcagtt taaactatca gtgtttgaca 15120
ggatatattg gcgggtaaac ctaagagaaa agagcgttta ttagaataat cggatattta 15180
aaagggcgtg aaaaggttta tccgttcgtc catttgtatg tgcatgcca ccaagggtt 15240

ccccagatct ggcgcgggcc agcgagacga gcaagattgg ccgcccgcgc aaacgatccg 15300
acagcgcgcc cagcacaggt gcgcaggcaa attgcaccaa cgcatacagc gccagcagaa 15360
tgccatagtg ggcggtgacg tcgttcgagt gaaccagatc gcgcaggagg cccggcagca 15420
ccggcataat caggccgatg ccgacagcgt cgagcgcgac agtgctcaga attacgatca 15480
ggggtatgtt gggtttcacg tctggcctcc ggaccagcct ccgctgggtcc gattgaacgc 15540
gcggattctt tatcactgat aagttggtgg acatattatg tttatcagtg ataaagtgtc 15600
aagcatgaca aagttgcagc cgaatacagt gatccgtgcc gccctggacc tgttgaacga 15660
ggtcggcgta gacggtctga cgacacgcaa actggcgga cggttggggg ttcagcagcc 15720
ggcgctttac tggcacttca ggaacaagcg ggcgctgctc gacgcactgg ccgaagccat 15780
gctggcggag aatcatacgc attcgggtgcc gagagccgac gacgactggc gctcatttct 15840
gatcggaat gcccgagct tcaggcaggc gctgctcgcc taccgcatg gcgcgcgcat 15900
ccatgccggc acgcgaccgg gcgcaccgca gatggaaacg gccgacgcgc agcttcgctt 15960
cctctgcgag gcgggttttt cggccgggga cgccgtcaat gcgctgatga caatcagcta 16020
cttactgtt ggggccgtgc ttgaggagca ggccggcgac agcgatgccg gcgagcgcgg 16080
cggcaccgtt gaacaggctc cgctctcgcc gctgttgcg gcgcgatag acgccttcga 16140
cgaagccggt ccggacgcag cgctcgagca gggactcgcg gtgattgtcg atggattggc 16200
gaaaaggagg ctcgttgtca ggaacgttga aggaccgaga aagggtgacg attgatcagg 16260
accgctgccg gagcgcaacc cactcactac agcagagcca tgtagacaac atccccccc 16320
cctttccacc gcgtcagacg ccgtagcag ccgctacgg gctttttcat gccctgccct 16380
agcgtccaag cctcacggcc gcgctcgcc tctctggcg ccttctggcg ctcttcgct 16440
tcctcgctca ctgactcgct gcgctcggtc gtccggctgc ggcgagcgg atcagctcac 16500

tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga 16560
gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 16620
aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 16680
ccgacaggac tataaagata ccaggcggtt cccctggaa gtcctcgt gcgctctcct 16740
gttccgacct tgccgcttac cggataacct tccgcctttc tcccttcggg aagcgtggcg 16800
cttttccgct gcataaccct gcttcggggt cattatagcg attttttcgg tataatccatc 16860
ctttttcgca cgatatacag gattttgcca aagggttcgt gtagactttc cttggtgtat 16920
ccaacggcgt cagccgggca ggataggtga agtaggcca cccgcgagcg ggtgttcctt 16980
cttactgtc ccttattcgc acctggcggt gctcaacggg aatcctgctc tgcgaggctg 17040
gccggctacc gccggcgtaa cagatgaggg caagcggatg gctgatgaaa ccaagccaac 17100
caggaagggc agcccaccta tcaagggtga ctgccttcca gacgaacgaa gagcgattga 17160
ggaaaaggcg gcggcgccg gcatgagcct gtcggcctac ctgctggccg tcggccaggg 17220
ctacaaaatc acgggcgtcg tggactatga gcacgtccgc gagctggccc gcatcaatgg 17280
cgacctgggc cgctgggcg gcctgctgaa actctggctc accgacgacc cgcgcacggc 17340
gcggttcggg gatgccacga tcctcgccct gctggcgaag atcgaagaga agcaggacga 17400
gcttggcaag gtcgatgagg gcgtggtccg cccgagggca gagccatgac ttttttagcc 17460
gctaaaacgg ccgggggggtg cgcgtgattg ccaagcacgt ccccatgcgc tccatcaaga 17520
agagcgactt cgcggagctg gtgaagtaca tcaccgacga gcaaggcaag accgagcgcc 17580
tttgcgacgc tca 17593

<210> 43

<211> 16954

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 43

ccgggctggt tgcctcgcc gctgggctgg cggcgtcta tggcctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt	480
ttttcgcca ccgtaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgct gaccgcgcac gccgaagggg ggtgcccccc	600

cttctcgaac cctcccggcc cgctaacgcg ggctcccat cccccaggg gctgcgccc 660

tcggccgcga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgcg cgcccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840

ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatata gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatactata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgctt aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgcccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttcgggag actgtcatat gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgcccggc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcattcaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggtttca 2100

aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220

aattagcttc ttggggatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgac ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gagggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggatcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggccccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaaacacgc ctgggtcaat gatgacctgg 4500

tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gtcagtatc 4620

gctcgggacg cacggcgcg cttacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gtcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggctg tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcattc catcctcggc gcacttaata tttcgctatt 5100

ctggagcttg ttgtttatct cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtccctgggg ctatttgctg aactgcgggc gtggcgctgt tgggtgttgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgctggcaa ctggcgggcg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcacaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgctt tacaacggct ctcccgctga cgccgtcccg gactgatggg ctgcctgtat 6060

cgagtggatga ttttgtgccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggctgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggatccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
aagaactgga ggggtgggtg caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320
tcctttgggt ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcg catgagggtt 7440
tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500
gcactattga tcatccgata gctctgcaa ggcgcttgca caatgcaagt caaacgttgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaatc 7920
tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980
cctagctgat tctggagtga ccagaggggt catgacttga gcctaaaatc cgcgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100
tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640

tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctacccgcg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcgctgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcggggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcttgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggctgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttggggcgag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggctac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgtcag gatctcgata agatacgttc atttgtcaa gcagcaaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatggt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaaccaga 10680

atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740

gtgtaagcgc ccactccaca tctccacteg acctgcaggc atgcaagctt gagattaaaa 10800

tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860

atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920

aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaaagt ttttttttat 10980

acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040

atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100

tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgtat ttactaatt 11160

cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220

ttgattacac tctaaatctc gatgacccga caaaaagag cgaacaaaga aataatcttg 11280

tgcattcgaa tatgatggaa gatTTTTTcc cccttattct aaatgttgac atagcgtgta 11340

tgttatataa acaaaaagaa attgtacaaa cttcttttcc ttctcttttt attttatctc 11400

tatgttgtgg atttggaatg ccctgatcgt ttctgttacc gtgattggca tggaagtgat 11460

tgctgcactg gcacacaaat acatcatgca cggctgggggt tggggatggc atctttcaca 11520

tcatgaaccg cgtaaagggt cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580

attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgctcc agtggattgg 11640

cgcagggtatg acggcgatatg gattactcta ttttatgggt caccacgggc tggatcatca 11700

acgttggcca ttccgctata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760

ccgtatgcat caccgcgtca ggggcaaaga aggttgtgtt tcttttggct tctctatgc 11820

gccgcccctg tcaaaacttc aggcgacgct ccgggaaaga catggcgcta gagcgggcgc 11880

tgccagagat gcgcagggcg gggaggatga gcccgcatcc gggaagtaag ggctgacca 11940

gaggcgcca gcagcagcgt taatttttcg ggcgtggcgt ttgactgccg ctgatcccaa 12000
agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gtcacaatt 12060
ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc 12120
taactcacat taattgcgtt gcgtcactg cccgctttcc agtcgggaaa cctgtcgtgc 12180
cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggccaaag 12240
acaaaagggc gacattcaac cgattgaggg agggaaagta aatattgacg gaaattattc 12300
attaaagggtg aattatcacc gtcaccgact tgagccattt gggaattaga gccagcaaaa 12360
tcaccagtag caccattacc attagcaagg ccggaacgt caccaatgaa accatcgata 12420
gcagcacctg aatcagtagc gacagaatca agtttgcctt tagcgtcaga ctgtagcgcg 12480
ttttcatcgg cattttcggg catagcccc ttattagcgt ttgccatctt ttcataatca 12540
aatcacccgg aaccagagcc accaccggaa ccgcctccct cagagccgcc accctcagaa 12600
ccgccaccct cagagccacc accctcagag ccgccaccag aaccaccacc agagccgccg 12660
ccagcattga caggaggccc gatctagtaa catagatgac accgcgcgcg ataatttattc 12720
ctagtttgcg cgctatattt tgttttctat cgcgtattaa atgtataatt gcgggactct 12780
aatcataaaa acccatctca taaataacgt catgcattac atgttaatta ttacatgctt 12840
aacgtaattc aacagaaatt atatgataat catcgcaaga ccggcaacag gattcaatct 12900
taagaaactt tattgcaaaa tgtttgaacg atcggggatc atccgggtct gtggcgggaa 12960
ctccacgaaa atatccgaac gcagcaagat atcgcgggtgc atctcgggtct tgcctgggca 13020
gtcgccgccg acgcggttga tgtggacgcc gggcccgatc atattgtcgc tcaggatcgt 13080
ggcgttgtgc ttgtcggccg ttgctgtcgt aatgatatcg gcaccttca ccgcctgttc 13140
cgcagagatc ccgtgggcga agaactccag catgagatcc ccgcgctgga ggatcatcca 13200

gccggcgctcc cggaaaacga ttccgaagcc caacctttca tagaaggcgg cgggtggaatc 13260
gaaatctcgt gatggcaggt tgggcgtcgc ttggtcggtc atttcgaacc ccagagtccc 13320
gctcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc gggagcggcg 13380
ataccgtaaa gcacgaggaa ggggtcagcc cattcgccgc caagctcttc agcaatatca 13440
cgggtagcca acgctatgtc ctgatagcgg tccgccacac ccagccggcc acagtcgatg 13500
aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc gccatgggtc 13560
acgacgagat catcgccgtc gggcatgccc gccttgagcc tggcgaacag ttcggctggc 13620
gcgagcccct gatgctcttc gtccagatca tctgatcga caagaccggc ttccatccga 13680
gtacgtgctc gctcgatgcg atgtttcgct tgggtggcga atgggcaggt agccggatca 13740
agcgtatgca gccgccgat tgcacagcc atgatggata ctttctcggc aggagcaagg 13800
tgagatgaca ggagatcctg ccccggaact tcgccaata gcagccagtc ccttcccgt 13860
tcagtgaaa cgtcgagcac agctgcgcaa ggaacgcccg tcgtggccag ccacgatagc 13920
cgcgctgcct cgtcctgcag ttcattcagg gcaccggaca ggtcggctctt gacaaaaaga 13980
accgggcgcc cctgcgctga cagccggaac acggcggcat cagagcagcc gattgtctgt 14040
tgtgccagt catagccgaa tagcctctcc acccaagcgg ccggagaacc tgcgtgcaat 14100
ccatcttggt caatcatgcg aaacgatcca gatccggtgc agattatttg gattgagagt 14160
gaatatgaga ctctaattgg ataccgaggg gaatttatgg aacgtcagtg gagcattttt 14220
gacaagaaat atttgctagc tgatagtgc cttaggcgac ttttgaacgc gcaataatgg 14280
tttctgacgt atgtgcttag ctcatataac tccagaaacc cgcggtgag tggctccttc 14340
aacgttgccg ttctgtcagt tccaaacgta aaacggcttg tcccgcgtca tcggcggggg 14400
tcataacgtg actcccttaa ttctccgctc atgatcagat tgcgttttc cgccttcagt 14460

ttaaactatc agtgtttgac aggatatatt ggcgggtaaa cctaagagaa aagagcgttt 14520
attagaataa tcggatattt aaaagggcgt gaaaaggttt atccgttcgt ccatttgtat 14580
gtgcatgcca accacagggt tccccagatc tggcgccggc cagcgagacg agcaagattg 14640
gccgccgccc gaaacgatcc gacagcgcg cagcacagg tgcgcaggca aattgcacca 14700
acgcatacag cgccagcaga atgccatagt gggcggtgac gtcgttcgag tgaaccagat 14760
cgcgcaggag gcccggcagc accggcataa tcaggccgat gccgacagcg tcgagcgcg 14820
cagtgtcag aattacgatc aggggtatgt tgggtttcac gtctggcctc cggaccagcc 14880
tccgctggtc cgattgaacg cgcggattct ttatcactga taagttggtg gacatattat 14940
gtttatcagt gataaagtgt caagcatgac aaagttgcag ccgaatacag tgatccgtgc 15000
cgccctggac ctgttgaacg aggtcggcgt agacggtctg acgacacgca aactggcgga 15060
acggttgggg gttcagcagc cggcgcttta ctggcacttc aggaacaagc gggcgctgct 15120
cgacgcactg gccgaagcca tgctggcgga gaatcacag cattcggtgc cgagagccga 15180
cgacgactgg cgctcatttc tgatcgggaa tgcccgcagc ttcaggcagg cgctgctcgc 15240
ctaccgcgat ggcgcgcgca tccatgccgg cacgcgaccg ggcgcaccgc agatggaaac 15300
ggccgacgcg cagcttcgct tcctctgcga ggcgggtttt tcggccgggg acgccgtcaa 15360
tgcgctgatg acaatcagct acttcactgt tggggccgtg cttgaggagc aggccggcga 15420
cagcgatgcc ggcgagcgcg gcggcaccgt tgaacaggct ccgctctcgc cgctgttgcg 15480
ggccgcgata gacgccttcg acgaagccgg tccggacgca gcgttcgagc agggactcgc 15540
ggtgattgtc gatggattgg cgaaaaggag gtcggtgtc aggaacgttg aaggaccgag 15600
aaagggtgac gattgatcag gaccgctgcc ggagcgcaac cactcacta cagcagagcc 15660
atgtagacaa catccctcc cctttccac cgcgtcagac gcccgtagca gcccgctacg 15720

ggcttttttca tgccctgccc tagcgtccaa gcctcacggc cgcgctcggc ctctctggcg 15780

gccttctggc gctcttccgc ttctctgctc actgactcgc tgcgctcggc cgttcggctg 15840

cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc tatccacaga atcaggggat 15900

aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 15960

gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc 16020

tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 16080

agctccctcg tgcgctctcc tgttccgacc ctgcgcgtta ccggatacct gtccgccttt 16140

ctcccttcgg gaagcgtggc gcttttccgc tgcataacc tgcctcgggg tcattatagc 16200

gattttttcg gtatatccat cctttttcgc acgatataca ggattttgcc aaagggttcg 16260

tgtagacttt ccttggtgta tccaacggcg tcagccgggc aggataggcg aagtagggcc 16320

accgcgagc ggggtgttct tcttactgt cccttattcg cacctggcgg tgctcaacgg 16380

gaatcctgct ctgcgaggct ggccggctac cgccggcgta acagatgagg gcaagcggat 16440

ggctgatgaa accaagccaa ccaggaaggc cagcccacct atcaagggtg actgccttcc 16500

agacgaacga agagcgattg aggaaaaggc ggcggcggcc ggcatgagcc tgcggccta 16560

cctgctggcc gtcggccagg gctacaaaat cacgggcgtc gtggactatg agcacgtccg 16620

cgagctggcc cgcataaatg gcgacctggc ccgcctgggc ggctgctga aactctggct 16680

caccgacgac ccgcgcacgg cgcggttcgg tgatgccacg atcctcgccc tgctggcgaa 16740

gatcgaagag aagcaggacg agcttgcaa ggtcatgatg ggcgtggtcc gcccgagggc 16800

agagccatga ctttttttagc cgctaaaacg gccggggggc gcgcgtgatt gccaagcacg 16860

tccccatgcg ctccatcaag aagagcgact tcgcggagct ggtgaagtac atcaccgacg 16920

agcaaggcaa gaccgagcgc ctttgcgacg ctca 16954

<210> 44
<211> 16954
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 44
ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcgttatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgtgta cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcggcca cgcctaacct gtcttttaac ctgcttttaa accaatatct ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaac cctcccggcc cgctaacgcg ggctcccat cccccaggg gctgcgcccc 660

tggccgcga acggcctcac cccaaaaatg gcagcgcctg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcatcagc ccgagcgcga cggccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtggagg cggcggcctg 840

ggtggcggcc tgcccttcac ttcggccgct ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttataa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatct gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaagtcttg aaaccagga caataacct atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgcgcctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatat gcgtaaaaca gccagcgcgt gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgcccggc 1740

tgtatgcgcg aggttaccga ctgcggcctg agtttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtccggcggg gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat tgtggtttca 2100

aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220

aattagcttc ttggggatc ttttaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgctc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttcgcgat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcggttcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgtgcc aaacgacacg gccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccgggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccagc 4320
ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380
gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggccttg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcattgcaa acgctagggc cttgtggggt cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860
catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctggtgga tgccatctt catcctcggc gcacttaata tttcgctatt 5100
ctggagcttg ttgtttattt cgggtotaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtccctggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgtgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgccctggcaa ctggcgggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc ttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcacaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgct tacaacggct ctcccgtga cgccgtccg gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatatgtggtg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300

aaatcctggt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggt 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcgggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgctctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctctttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggetgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa ggcgcttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaato 7920

tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggg catgacttga gcctaaaatc cgccgcctcc 8040

accattttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctcttg cggtcttgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaatctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggcccc 8640

tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggctgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctacccgc 8940

ttgagcagac atcaccatgc ctgaaatcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctggtgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccatttcgg 9360

accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcttgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccc a gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500

tcttcttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttccctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacagggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctctgttact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt agagataaaa 10800

taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac atacacgcta 10860

tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc acaagattat 10920

ttctttgttc gctctttttg gtcgggtcat cgagatttag agtgtaatca aagatactgt 10980

catctogaga gcgttgcaaca ggctgctgtt tgccaaattg gatgtttgcc gaattagtaa 11040

aatacgcaag catttcttac ctttccgctc ctttttcta attctccaa agactaaatg 11100

aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga tataaacttg 11160

cagcacgtaa gaccaaagca aattggtaac tattcttgtg taaaaacatg tataaaaaaa 11220

aacttttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat tgcttgata 11280

tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta tttataaaaa 11340

gaagaatgtg ccatgcttcc gaatttcttt tcactttctt ttccttatct attttaatct 11400

catgttgtgg atttgaatg cctgatcgt tttcgttacc gtgattggca tggaagtgat 11460

tgctgcactg gcacacaaat acatcatgca cggctggggt tggggatggc atctttcaca 11520

tcatgaaccg cgtaaagggt cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580

attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgctcc agtggattgg 11640

cgcagggtatg acggcgtatg gattactcta ttttatgggt caccacgggc tgggtgcatca 11700

acgttggcca ttccgctata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760

ccgtatgcat caccgcgtca ggggcaaaga aggttgtgtt tcttttggct tcctctatgc 11820

gccgcccctg tcaaaacttc aggcgacgct ccgggaaaga catggcgcta gagcggggcg 11880

tgccagagat gcgcagggcg gggaggatga gcccgcattc gggaagtaag ggcctgacca 11940

gaggcgggcca gcagcagcgt taatttttcg ggcgtggctg ttgactgccg ctgatcccaa 12000

agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gtcacaatt 12060

ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc 12120

taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc 12180

cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tggggcaaag 12240

acaaaagggc gacattcaac cgattgaggg aggggaaggta aatattgacg gaaattattc 12300

attaaagggtg aattatcacc gtcacgcact tgagccattt gggaattaga gccagcaaaa 12360

tcaccagtag caccattacc attagcaagg ccggaaacgt caccaatgaa accatcgata 12420

gcagcaccgt aatcagtagc gacagaatca agtttgcctt tagcgtcaga ctgtagcgcg 12480

ttttcatcgg cattttcggt catagcccc ttattagcgt ttgccatctt ttcataatca 12540

aatcaccgg aaccagagcc accaccggaa ccgcctccct cagagccgcc accctcagaa 12600

ccgccaccct cagagccacc accctcagag ccgccaccag aaccaccacc agagccgccc 12660

ccagcattga caggaggccc gatctagtaa catagatgac accgcgcgcg ataatttatc 12720

ctagtttgcg cgctatattt tgttttctat cgcgtattaa atgtataatt gcgggactct 12780

aatcataaaa acccatctca taaataacgt catgcattac atgttaatta ttacatgctt 12840

aacgtaattc aacagaaatt atatgataat catcgcaaga ccggcaacag gattcaatct 12900

taagaaactt tattgcaaaa tgtttgaacg atcggggatc atccgggtct gtggcgggaa 12960

ctccacgaaa atatccgaac gcagcaagat atcgcggtgc atctcggtct tgcctgggca 13020

gtcgcgcgcg acgccgttga tgtggacgcc gggcccgatc atattgtcgc tcaggatcgt 13080

ggcgttgtgc ttgtcggccg ttgtgtcgt aatgatatcg gcaccttcga ccgcctgttc 13140
cgagagatc ccgtgggcca agaactccag catgagatcc ccgcgtgga ggatcatcca 13200
gccggcgtcc cggaaaacga ttccgaagcc caacctttca tagaaggcgg cggtggaatc 13260
gaaatctcgt gatggcaggt tgggcgtcgc ttggtcggtc atttcgaacc ccagagtccc 13320
gctcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc gggagcggcg 13380
ataccgtaaa gcacaggaa gcggtcagcc cattcgccgc caagctcttc agcaatatca 13440
cgggtagcca acgctatgtc ctgatagcgg tccgccacac ccagccggcc acagtcatg 13500
aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc gccatgggtc 13560
acgacgagat catcgccgtc gggcatgcgc gccttgagcc tggcgaacag ttcggtggc 13620
gogagcccct gatgctcttc gtccagatca tcctgatoga caagaccggc ttccatccga 13680
gtacgtgtc gctcgatgcg atgtttcgtc tggtggtcga atgggcaggt agccggatca 13740
agcgtatgca gccgccgat tgcacagcc atgatggata ctttctcggc aggagcaagg 13800
tgagatgaca ggagatcctg ccccgccact tcgccaata gcagccagtc ctttcccgct 13860
tcagtgaaa cgtcgagcac agctgcgcaa ggaacgcccg tcgtggccag ccacgatagc 13920
cgcgctgcct cgtcctgcag ttcattcagg gcaccggaca ggtcggtctt gacaaaaaga 13980
accggggccc cctgcgctga cagccggaac acggcggcat cagagcagcc gattgtctgt 14040
tgtgcccagt catagccgaa tagcctctcc acccaagcgg ccggagaacc tgcgtgcaat 14100
ccatcttggt caatcatgcg aaacgatcca gatccggtgc agattatttg gattgagagt 14160
gaatatgaga ctctaattgg ataccgagg gaatttatgg aacgtcagt gagcattttt 14220
gacaagaaat atttgctagc tgatagtac cttaggcgac ttttgaacgc gcaataatgg 14280
tttctgacgt atgtgcttag ctcatataac tccagaaacc cggcgctgag tggctccttc 14340

aacgttgcgg ttctgtcagt tccaaacgta aaacggcttg tcccgcgtca tcggcggggg 14400
tcataacgtg actcccttaa ttctccgctc atgatcagat tgtcgtttcc cgccttcagt 14460
ttaaactatc agtgtttgac aggatatatt ggcggtataa cctaagagaa aagagcgttt 14520
attagaataa tcggatattt aaaagggcgt gaaaaggttt atccgttcgt ccatttgtat 14580
gtgcatgcca accacagggg tccccagatc tggcgccggc cagcgagacg agcaagattg 14640
gccgcgccc gaaacgatcc gacagcgcgc ccagcacagg tgcgcaggca aattgcacca 14700
acgcatacag cgccagcaga atgccatagt gggcggtgac gtcgttcgag tgaaccagat 14760
cgcgaggag gcccggcagc accggcataa tcaggccgat gccgacagcg tcgagcgcca 14820
cagtgtcag aattacgatc aggggtatgt tgggtttcac gtctggcctc cggaccagcc 14880
tccgtggtc cgattgaacg cgcggttct ttatcactga taagttggtg gacatattat 14940
gtttatcagt gataaagtgt caagcatgac aaagttgcag ccgaatacag tgatccgtgc 15000
cgccctggac ctgttgaacg aggtcggcgt agacggtctg acgacacgca aactggcgga 15060
acggttgggg gttcagcagc cggcgcttta ctggcacttc aggaacaagc gggcgctgct 15120
cgacgcactg gccgaagcca tgctggcgga gaatcatac cattcggtgc cgagagccga 15180
cgacgactgg cgctcatttc tgatcgggaa tgcccgcagc ttcaggcagg cgctgctcgc 15240
ctaccgcgat ggcgcgcgca tccatgccgg cagcgaccg ggcgaccgc agatggaaac 15300
ggccgacgcg cagcttcgct tcctctgcga ggcgggtttt tcggccgggg acgccgtcaa 15360
tgcgctgatg acaatcagct acttcaactgt tggggccgtg cttgaggagc aggcggcgga 15420
cagcgatgcc ggcgagcgcg gcggcaccgt tgaacaggct ccgctctcgc cgctgttgcg 15480
ggccgcgata gacgccttcg acgaagccgg tccggacgca gcgttcgagc agggactcgc 15540
ggtgattgtc gatggattgg cgaaaaggag gtcgttgtc aggaacgttg aaggaccgag 15600

aaaggggtgac gattgatcag gaccgctgcc ggagcgcaac ccactcacta cagcagagcc 15660
atgtagacaa catccccctcc ccctttccac cgcgtcagac gcccgtagca gcccgctacg 15720
ggctttttca tgcctgccc tagcgtccaa gcctcacggc cgcgctcggc ctctctggcg 15780
gccttctggc gctcttcgc ttctcgtc actgactcgc tgcgctcggc cgttcggctg 15840
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc tatccacaga atcaggggat 15900
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 15960
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc 16020
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt tccccctgga 16080
agctccctcg tgcgctctcc tgttcgacc ctgccgctta ccggatacct gtccgccttt 16140
ctcccttcgg gaagcgtggc gcttttcgc tgcataacc tgcttcgggg tcattatagc 16200
gattttttcg gtatatccat ctttttcgc acgatataca ggattttgcc aaagggttcg 16260
tgtagacttt ctttggtgta tccaacggcg tcagccgggc aggataggcg aagtagggcc 16320
accgcgagc ggtgttcct tcttactgt cccttattcg cacctggcgg tgctcaacgg 16380
gaatcctgct ctgcgaggct ggccggctac cgccggcgta acagatgagg gcaagcggat 16440
ggctgatgaa accaagccaa ccaggaaggc cagcccacct atcaagggtg actgccttcc 16500
agacgaacga agagcgattg aggaaaaggc ggcggcggcc ggcatgagcc tgcggccta 16560
cctgctggcc gtcggccagg gctacaaaat cagggcgctc gtggactatg agcacgtccg 16620
cgagctggcc cgcacaaatg gcgacctggc ccgcctgggc ggctgctga aactctggct 16680
caccgacgac ccgcgcacgg cgcggttcgg tgatgccacg atcctcgccc tgctggcgaa 16740
gatcgaagag aagcaggacg agcttggcaa ggtcatgatg ggcgtggctc gcccgagggc 16800
agagccatga cttttttagc cgctaaaacg gccggggggt gcgcgtgatt gccaaagcac 16860

tcccatgcg ctccatcaag aagagcgact tcgcgagct ggtgaagtac atcaccgacg 16920

agcaaggcaa gaccgagcgc ctttgcgacg ctca 16954

<210> 45

<211> 19491

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (18970)..(18970)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (19178)..(19178)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (19269)..(19269)

<223> n is a, c, g, or t

<400> 45

agcttggtac cgagctcgga tccactagta acggcgcga gtgtgctgga attcgccctt 60

gacggccagt gaattcgagc tcggtaccgc gggatctttc gacactgaaa tacgtcgagc 120

ctgctccgct tggaagcggc gaggagcctc gtcctgtcac aactaccaac atggagtacg 180

ataagggccca gttccgccag ctcatthaaga gccagttcat gggcgttggc atgatggccg 240

tcatgcatct gtacttcaag tacaccaacg ctcttctgat ccagtcgatc atccgctgaa 300

ggcgctttcg aatctgggta agatccacgt cttegggaag ccagcgactg gtgacctcca 360

gcgtcccttt aaggctgcc aagctttct cagccagggc cagccaaga ccgacaaggc 420

ctccctccag aacgccgaga agaactggag ggggtggtgc aaggaggagt aagctcctta 480

ttgaagtcgg aggacggagc ggtgtcaaga ggatattctt cgactctgta ttatagataa 540

gatgatgagg aattggaggt agcatagctt catttggatt tgctttccag gctgagactc 600

tagcttggag catagagggg cctttggctt tcaatattct caagtatctc gagtttgaac 660

ttattccctg tgaacctttt attcaccaat gagcattgga atgaacatga atctgaggac 720

tgcaatcgcc atgagggttt cgaaatacat ccggatgtcg aaggcttggg gcacctgcgt 780

tggttgaatt tagaacgtgg cactattgat catccgatag ctctgcaaag ggcgttgac 840

aatgcaagtc aaacgttgct agcagttcca ggtggaatgt tatgatgagc attgtattaa 900

atcaggagat atagcatgat ctctagttag ctcaccacaa aagtcagacg gcgtaaccaa 960

aagtcacaca acacaagctg taaggatttc ggcacggcta cggaagacgg agaagccacc 1020

ttcagtggac tcgagtacca tttaattcta tttgtgtttg atcgagacct aatacagccc 1080

ctacaacgac catcaaagtc gtatagctac cagtgaggaa gtggactcaa atcgacttca 1140

gcaacatctc ctggataaac tttaagccta aactatacag aataagatag gtggagagct 1200

tataccgagc tcccaaattc gtccagatca tgggtgaccg gtgcctggat ctccctatag 1260

aatcatcctt attcgttgac ctagctgatt ctggagtgc ccagagggtc atgacttgag 1320

cctaaaatcc gccgcctcca ccattttagt aaaaatgtga cgaactcgtg agctctgtac 1380

agtgaccggg gactctttct ggcattgcga gagacggacg gacgcagaga gaagggctga 1440

gtaataagcc actggccaga cagctctggc ggctctgagg tgcagtggat gattattaat 1500

ccgggaccgg ccgcccctcc gccccgaagt ggaaaggctg gtgtgcccct cgttgaccaa 1560

gaatctattg catcatcgga gaatatggag cttcatcgaa tcaccggcag taagcgaagg 1620

agaatgtgaa gccaggggtg tatagccgtc ggcgaaatag catgccatta acctaggtac 1680

agaagtccaa ttgcttccga tctggtaaaa gattcacgag atagtacctt ctccgaagta 1740

ggtagagcga gtacccggcg cgtaagctcc ctaattggcc catccggcat ctgtagggcg 1800

tccaaatata gtgcctctcc tgctttgccc ggtgtatgaa accggaaagg ccgctcagga 1860

gctggccagc ggcgcagacc gggaacacaa gctggcagtc gacctatccg gtgctctgca 1920

ctcgacctgc tgaggctccct cagtcctctg taggcagctt tgccccgtct gtccgccccg 1980

tgtgtcggcg ggggtgacaa ggtcgttgcg tcagtccaac atttggtgcc atattttcct 2040

gctctcccca ccagctgctc ttttcttttc tctttctttt cccatcttca gtatattcat 2100

cttcccatcc aagaaccttt atttcccta agtaagtact ttgtacatc catactccat 2160

ccttcccatc ccttattcct ttgaaccttt cagttcgagc tttcccactt catcgagct 2220

tgactaacag ctaccccgct tgagcagaca tcaccatgct gtcgaagctg cagtcaatca 2280

gcgtcaaggc ccgccgctt gaactagccc gcgacatcac gcggccaaa gtctgcctgc 2340

atgctcagcg gtgctcgta gttcggctgc gagtggcagc accacagaca gaggaggcg 2400

tgggaaccgt gcaggctgcc ggcgggggcg atgagcacag cgccgatgta gcactccagc 2460

agcttgaccg ggctatcgca gagcgtcgtg cccggcgcaa acgggagcag ctgtcatacc 2520

aggctgccgc cattgcagca tcaattggcg tgtcaggcat tgccatcttc gccacctacc 2580

tgagatttgc catgcacatg accgtggggc gcgcagtgcc atggggtgaa gtggctggca 2640

ctctcctctt ggtgggttgt ggcgcgctcg gcatggagat gtatgcccgc tatgcacaca 2700

aagccatctg gcatgagtcg cctctgggct ggctgctgca caagagccac cacacacctc 2760

gcactggacc ctttgaagcc aacgacttgt ttgcaatcat caatggactg cccgccatgc 2820

tcctgtgtac ctttggcttc tggctgcca acgtcctggg ggcggcctgc tttggagcgg 2880

ggctgggcat cacgctatac ggcatggcat atatgtttgt acacgatggc ctggtgcaca 2940
ggcgctttcc caccggggccc atcgctggcc tgccctacat gaagcgctg acagtggccc 3000
accagctaca ccacagcggc aagtacggtg gcgcgccctg gggatatgtc ttgggtccac 3060
aggagctgca gcacattcca ggtgcggcgg aggaggtgga gcgactggtc ctggaactgg 3120
actggtccaa gcggtagggt gcggaaccag gcacgctggt ttcacacctc atgcctgtga 3180
taaggtgtgg ctagagcgat gcgtgtgaga cgggtatgtc acggtcgact ggtctgatgg 3240
ccaatggcat cggccatgtc tggatcacac gggctggttg cctgggtgaa ggtgatgcac 3300
atcatcatgt gcggttgag gggctggcac agtgtgggct gaactggagc agttgtccag 3360
gctggcggtt aatcagtgag ggtttgtgat tggcggttga gaagcaatga ctccgcccat 3420
attctatttg tgggagctga gatgatggca tgcttgggat gtgcatggat catggtagtg 3480
cagcaaaacta tattcaccta gggctgttgg taggatcagg tgaggccttg cacattgcat 3540
gatgtactcg tcatggtgtg ttggtgagag gatggatgtg gatggatgtg tattctcaga 3600
cgtagacctt gactggaggc ttgatcgaga gagtgggccg tattctttga gaggggaggc 3660
tcgtgccaga aatggtgagt ggatgactgt gacgctgtac attgcaggca ggtgagatgc 3720
actgtctcga ttgtaaaata cattcagatg caagcttggc gtaatcatgg tcatagctgt 3780
ttcctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc ggaagcataa 3840
agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg ttgcgctcac 3900
tgcccgcctt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 3960
cggggagagg cggtttgcgt attgggcca agacaaaagg gcgacattca accgattgag 4020
ggaggggaagg taaatattga cggaaattat tcattaaagg tgaattatca ccgtcaccga 4080
cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa 4140

ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta gcgacagaat 4200

caagtttgcc tttagcgta gactgtagcg cgttttcatc ggcattttcg gtcatagccc 4260

ccttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag ccaccaccgg 4320

aaccgcctcc ctgagagccg ccaccctcag aaccgccacc ctgagagcca ccaccctcag 4380

agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggg ccgatctagt 4440

aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat tttgttttct 4500

atcgcgatt aaatgtataa ttgcgggact ctaatcataa aaacccatct cataaataac 4560

gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata 4620

atcatcgcaa gaccggcaac aggattcaat cttaagaaac tttattgcca aatgtttgaa 4680

cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga acgcagcaag 4740

atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt gatgtggacg 4800

ccgggcgcga tcatattgtc gctcaggatc gtggcgttgt gcttgtcggc cgttgctgtc 4860

* gtaatgatat cggcacctc gaccgcctgt tccgcagaga tcccgtaggc gaagaactcc 4920

agcatgagat ccccgcgctg gaggatcatc cagccggcgt cccggaaaac gattccgaag 4980

cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag gttgggcgtc 5040

gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca agaaggcgat 5100

agaaggcgat gcgtgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag 5160

cccattcgcc gccaaagctct tcagcaatat caccggtagc caacgctatg tcctgatagc 5220

ggtcogccac acccagccgg ccacagtcga tgaatccaga aaagcggcca tttccacca 5280

tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg tcgggcatgc 5340

gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct tcgtccagat 5400

catcctgata gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg 5460

cttgggtggc gaatgggcag gtagccggat caagcgtatg cagccgccgc attgcatcag 5520

ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc tgccccggca 5580

cttcgcccaa tagcagccag tcccttcccg cttcagtgc aacgtcgagc acagctgcgc 5640

aaggaacgcc cgtcgtggcc agccaagata gccgcgctgc ctgcctcctgc agttcattca 5700

gggcaccgga caggtcggtc ttgacaaaaa gaaccggggc cccctgcgct gacagccgga 5760

acacggcggc atcagagcag ccgattgtct gttgtgcca gtcatagccg aatagcctct 5820

ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgatc 5880

cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag 5940

gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg 6000

accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt agctcattaa 6060

actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca gttccaaacg 6120

taaaacggct tgtcccgct catcggcggg ggtcataacg tgactccctt aattctccgc 6180

tcatgatcag attgtcggtt cccgccttca gtttaaaacta tcagtgtttg acaggatata 6240

ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcgatat taaaagggc 6300

gtgaaaaggc ttatccgttc gtccatttgt atgtgcatgc caaccacagg gttccccaga 6360

tctggcgccg gccagcgaga cgagcaagat tggccgccgc ccgaaacgat ccgacagcgc 6420

gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca gaatgccata 6480

gtgggcggtg acgtcgttcg agtgaaccag atcgcgcagg agggccggca gcaccggcat 6540

aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga tcaggggtat 6600

gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa cgcgcggtat 6660

ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg 6720

acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa cgaggtcggc 6780

gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca gccggcgctt 6840

tactggcact tcaggaacaa gggggcgctg ctcgacgcac tggccgaagc catgctggcg 6900

gagaatcata cgcatcgggt gccgagagcc gacgacgact ggcgctcatt tctgatcggg 6960

aatgcccgca gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg catccatgcc 7020

ggcacgcgac cgggcgccacc gcagatggaa acggccgacg cgcagcttcg cttcctctgc 7080

gaggcggggtt tttcggccgg ggacgccgtc aatgcgctga tgacaatcag ctacttcact 7140

gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg cggcggcacc 7200

gttgaacagg ctccgctctc gccgctgttg cgggcccga tagacgcctt cgacgaagcc 7260

ggtccggacg cagcgctcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg 7320

aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc aggaccgctg 7380

ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct ccccctttcc 7440

accgcgtcag acgcccgtag cagcccgtc cgggcttttt catgccctgc cctagcgctc 7500

aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc gcttctctgc 7560

tcactgactc gctgcgctcg gtcgttcggc tgccggcgagc ggtatcagct cactcaaagg 7620

cggtaatcag gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 7680

gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 7740

gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagac 7800

gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 7860

ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcggtg gcgcttttcc 7920

gctgcataac cctgcttcgg ggtcattata gcgatttttt cggatatatcc atccttttttc 7980

gcacgatata caggattttg ccaaaggggt cgtgtagact ttccttggtg tatccaacgg 8040

cgtcagccgg gcaggatagg tgaagtaggc ccacccgcga gcgggtgttc cttcttcact 8100

gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg ctggccggct 8160

accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag 8220

ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat tgaggaaaag 8280

gcggcgggcg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca gggctacaaa 8340

atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg ccgcatcaa tggcgacctg 8400

ggccgcctgg gcggcctgct gaaactctgg ctaccgcagc acccgcgac gcgcgggttc 8460

ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga cgagcttggc 8520

aaggatcatga tgggcgtggt ccgcccaggg gcagagccat gactttttta gccgctaaaa 8580

cgccgggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca agaagagcga 8640

cttcgaggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc gcctttgcga 8700

cgctcacccg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc cctgcaaacg 8760

cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt tgtggatacc 8820

tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact tgagggggcg 8880

actcaccgg cgcggcgttg acagatgagg ggcaggctcg atttcggccg gcgacgtgga 8940

gctggccagc ctgcgaaatc ggcgaaaacg cctgatttta cgcgagtttc ccacagatga 9000

tgtggacaag cctgggggata agtgccctgc ggtattgaca cttgaggggc gcgactactg 9060

acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga tgagggggcg 9120

acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc aagggtttcc 9180

gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca atatttataa 9240
accttgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacgccg aaggggggtg 9300
cccccccttc tcgaaccctc cgggcccgct aacgcgggcc tcccatcccc ccaggggctg 9360
cgcccctcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt ccttgccatt 9420
gccgggatcg gggcagtaac gggatgggcg atcagcccga gcgcgacgcc cggaagcatt 9480
gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag tgagggcggc 9540
ggcctgggtg ggggcctgcc cttcacttcg gcgctggggg cattcacgga cttcatggcg 9600
gggccggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc cgtgctcgtg 9660
ttcgggggtg cgataaacc agcgaaccat ttgaggtgat aggtaagatt ataccgaggt 9720
atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat ttaaaaagct 9780
accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat attgacaata 9840
ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga tttcaggggg 9900
caaggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca taaaaacttg 9960
catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt ctatcataat 10020
tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc gatgactttg 10080
tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg tgccaggtgc 10140
tgctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct gattacgtgc 10200
agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca tatcaccag 10260
tcaaaggggtg acagcaggct cataagacgc ccagcgtcg ccatagtgcg ttcaccgaat 10320
acgtgcgcaa caaccgtctt ccggagactg tcatacgcg aaaacagcca gcgctggcgc 10380
gatttagccc cgacatagcc ccactgttcg tccatttcgg cgcagacgat gacgtcactg 10440

ccccgctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga cgtaaaatcg 10500
tggttgaggcc aacgcccata atgcgggctg ttgcccggca tccaacgcca ttcattggcca 10560
tatcaatgat tttctggtgc gtaccggggt gagaagcggg gtaagtgaac tgcagttgcc 10620
atgtttttacg gcagtgagag cagagatagc gctgatgtcc ggcggtgctt ttgccgttac 10680
gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa gccactggag 10740
cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc cataattgtg 10800
gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac aactttgaaa 10860
aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg gagttcgtct 10920
tggtataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaaataata 10980
aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt 11040
aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctgggtgggag aaaatgaaaa 11100
cctatattta aaaatgacgg acagccggta taaagggacc acctatgatg tggaacggga 11160
aaaggacatg atgctatggc tggaaggaaa gctgcctggt ccaaagggtc tgcactttga 11220
acggcatgat ggctggagca atctgtcat gactgaggcc gatggcgctc tttgctcgga 11280
agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg agtgcatcag 11340
gctctttcac tccatcgaca tatcgattg tccctatacg aatagcttag acagccgctt 11400
agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg aaaactggga 11460
agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaaaagcc 11520
cgaagaggaa cttgtctttt ccacggcga cctgggagac agcaacatct ttgtgaaaga 11580
tggtcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca agtgggtatga 11640
cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt atgtcgagct 11700

atatttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt atattttact 11760
ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca 11820
ccgacttctt cgcacatcaag tgttttggct ctcaggccga ggcccacggc aagtatttgg 11880
gcaagggggtc gctggtattc gtgcagggca agattcggaa taccaagtac gagaaggacg 11940
gccagacggc ctacgggacc gacttcattg ccgataaggt ggattatctg gacaccaagg 12000
caccaggcgg gtcaaatcag gaataagggc acattgcccc ggcgtgagtc ggggcaatcc 12060
cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa gaactgatcg 12120
acgcgggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc atgcgtgcgc 12180
cccgcgaaac cttccagtcc gtcggctcga tgggccagca agctacggcc aagatcgagc 12240
gcgacagcgt gcaactggct cccctgccc tgcccgcc atcgccgcc gtggagcgtt 12300
cgcgctcgtc cgaacaggag gcggcagggt tggcgaagtc gatgaccatc gacacgcgag 12360
gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa caggtcagcg 12420
aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa atgcagcttt 12480
ccttgttcga tattgcgccg tggccggaca cgatgcgagc gatgccaaac gacacggccc 12540
gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg caaaacaagg 12600
tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgctcgag ctgcggggccg 12660
acgatgacga actggtgtgg cagcagggtg tggagtacgc gaagcgcacc cctatcggcg 12720
agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggtcg atcaatggcc 12780
ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg atgggcttca 12840
cgtccgaccg cgttgggcac ctggaatcgg tgctcgtgct gcaccgcttc cgcgtcctgg 12900
accgtggcaa gaaaacgtcc cgttgccagg tctgatcga cgaggaaatc gtcgtgctgt 12960

ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg tcgccgacgg 13020

cccgcaggat gttcgactat ttcagctcgc accgggagcc gtacccgctc aagctggaaa 13080

ccttccgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc gagcaggctc 13140

gcgaagcctg cgaagagttg cgaggcagcg gcctgggtgga acacgcctgg gtcaatgatg 13200

acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg gggttcagcag 13260

ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact tgcttcgctc 13320

agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag gattaaaatt 13380

gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc aggatttccg 13440

cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg tttacgagca 13500

cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg tggcattcgg 13560

cgcctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg acggccccaa 13620

ggacgctcac aaggcgcac tgtccggcgt tttcgtggag cccgaacagc gaggccgagg 13680

ggtcgcgggt atgctgctgc gggcgttgcc ggcgggttta ttgctcgtga tgatcgctcc 13740

acagattcca acgggaatct ggtggatgcy catcttcac ctcggcgcac ttaatatttc 13800

gctattctgg agcttgttgt ttatttcggt ctaccgcctg ccgggcgggg tcgcggcgac 13860

ggtaggcgct gtgcagccgc tgatggctgt gttcatctct gccgctctgc taggtagccc 13920

gatacgattg atggcgggtcc tgggggctat ttgcggaact gcgggcgtgg cgctgttggt 13980

gttgacacca aacgcagcgc tagatcctgt cggcgctcga gcgggcctgg cgggggcggg 14040

ttccatggcg ttcggaaccg tgctgaccgc caagtggcaa cctcccgtgc ctctgctcac 14100

ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag ctttagtggt 14160

tgatccgcca atcccgatgc ctacaggaac caatgttctc ggctggcgt ggctcggcct 14220

gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac tcgaacctac 14280

agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc cggggatgca 14340

tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcgggtgag caatggatag 14400

gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc ttcttcagcg 14460

gctttatcca gcgatttctt attatgtcgg catagtcttc aagatcgaca gcctgtcacg 14520

gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg agatgatatt 14580

tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct ccgcgagatc 14640

atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc ggtaacatga 14700

gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact gatgggctgc 14760

ctgtatcgag tggtgatttt gtgccgagct gccggtcggg gagctggttg ctggctggtg 14820

gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac acattgcgga 14880

cgtttttaat gtactggggt ggtttttctt ttcaccagtg agacgggcaa cagctgattg 14940

cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggg ttgccccagc 15000

aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat aaatcaaaag 15060

aatagcccgga gatagggttg agtggtgttc cagtttgga caagagtcca ctattaaaga 15120

acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg 15180

aaccatcacc caaatcaagt tttttggggt cgagggtgcc taaagcacta aatcggaacc 15240

ctaaaggag cccccgattt agagcttgac ggggaaagcc ggcgaaacgtg gcgagaaagg 15300

aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg ggaagggcga 15360

tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 15420

ttaagttggg taacgccagg gttttccag tcacgacgtt gtaaaacgac ggccagtga 15480

ttcgagctcg gtacccgggg atctttcgac actgaaatac gtcgagcctg ctccgcttgg 15540
aagcggcgag gagcctcgtc ctgtcacaac taccaacatg gagtacgata agggccagtt 15600
ccgccagctc attaagagcc agttcatggg cgttggcatg atggccgtca tgcattctgta 15660
cttcaagtac accaacgctc ttctgatcca gtcgatcatc cgctgaaggc gctttcgaat 15720
ctgggttaaga tccacgtctt cgggaagcca gcgactggtg acctccagcg tccctttaag 15780
gctgccaaca gctttctcag ccagggccag cccaagaccg acaaggcctc cctccagaac 15840
gccgagaaga actggagggg tgggtgtcaag gaggagtaag ctccttattg aagtoggagg 15900
acggagcggg gtcaagagga tattcttcga ctctgtatta tagataagat gatgaggaat 15960
tggaggtagc atagcttcat ttggatttgc tttccaggct gagactctag cttggagcat 16020
agagggtcct ttggctttca atattctcaa gtatctcgag tttgaactta ttccctgtga 16080
accttttatt caccaatgag cattggaatg aacatgaatc tgaggactgc aatcgccatg 16140
aggttttcga aatacatccg gatgtcgaag gcttggggca cctgcgttgg ttgaatttag 16200
aacgtggcac tattgatcat ccgatagctc tgcaaagggc gttgcacaat gcaagtcaaa 16260
cgttgctagc agttccaggg ggaatgttat gatgagcatt gtattaaatc aggagatata 16320
gcatgatctc tagttagctc accacaaaag tcagacggcg taaccaaag tcacacaaca 16380
caagctgtaa ggatttcggc acggctacgg aagacggaga agccacctc agtggactcg 16440
agtaccattt aattctattt gtgtttgatc gagacctaat acagccccta caacgaccat 16500
caaagtcgta tagctaccag tgaggaagtg gactcaaatc gacttcagca acatctcctg 16560
gataaacttt aagcctaaac tatacagaat aagatagggtg gagagcttat accgagctcc 16620
caaactctgc cagatcatgg ttgaccggtg cctggatctt cctatagaat catccttatt 16680
cgttgacctc gctgattctg gaggaccca gagggtcatg acttgagcct aaaatccgcc 16740

gcctccacca ttgttagaaa aatgtgacga actcgtgagc tctgtacagt gaccggtgac 16800
tctttctggc atgcggagag acggacggac gcagagagaa gggctgagta ataagccact 16860
ggccagacag ctctggcggc tctgaggtgc agtggatgat tattaatccg ggaccggccg 16920
ccccccgcc ccgaagtga aaggctggtg tgccccctgt tgaccaagaa tctattgcat 16980
catcggagaa tatggagctt catcgaatca ccggcagtaa gcgaaggaga atgtgaagcc 17040
aggggtgtat agccgtcggc gaaatagcat gccattaacc taggtacaga agtccaattg 17100
cttcgatctt ggtaaaagat tcacgagata gtaccttctc cgaagtaggt agagcgagta 17160
cccgccggt aagctcccta attggcccat ccggcatctg tagggcgctc aaatatcgtg 17220
cctctcctgc ttgtcccggt gtatgaaacc ggaaaggccg ctcaggagct ggccagcggc 17280
gcagaccggg aacacaagct ggcagtcgac ccatccggtg ctctgcactc gacctgctga 17340
ggtccctcag tccctggtag gcagctttgc ccgtctgtc cggccggtgt gtcggcgggg 17400
ttgacaaggc cgttgctgca gtccaacatt tgttgccata ttttcctgct ctccccacca 17460
gctgctcttt tcttttctct ttcttttccc atcttcagta tattcatctt cccatccaag 17520
aacctttatt tcccctaagt aagtactttg ctacatocat actccatcct tcccatccct 17580
tattcctttg aacctttcag ttcgagcttt ccacttcat cgcagcttga ctaacagcta 17640
ccccgcttga gcagacatca ccatgcctga actcaccgcg acgtctgtcg agaagtttct 17700
gatcgaaaag ttcgacagcg tctccgacct gatgcagctc tcggaggggc aagaatctcg 17760
tgctttcagc ttgatgtag gagggcgtgg atatgtcctg cgggtaaata gctgcgccga 17820
tggtttctac aaagatcgtt atgtttatcg gcactttgca tcggccgcgc tcccgatcc 17880
ggaagtgctt gacattgggg aattcagcga gagcctgacc tattgcatct cccgccgtgc 17940
acagggtgtc acgttgcaag acctgcctga aaccgaactg cccgctgttc tgcagccggt 18000

cgcgaggagcc atggatgcga tcgctgcggc cgatcttagc cagacgagcg gggtcgccc 18060
attcggaccg caaggaatcg gtcaatacac tacatggcgt gatttcatat gcgcgattgc 18120
tgatcccat gtgtatcact ggcaaactgt gatggacgac accgtcagtg cgtccgtcgc 18180
gcaggctctc gatgagctga tgctttgggc cgaggactgc cccgaagtcc ggcacctcgt 18240
gcacgcggat ttcggctcca acaatgtcct gacggacaat ggccgcataa cagcggtcac 18300
tgactggagc gaggcgatgt tcggggattc ccaatacgag gtcgccaaca tcttctctg 18360
gagggcgtgg ttggcttgta tggagcagca gacgcgctac ttcgagcgga ggcacccgga 18420
gcttgacagga tcgccgcggc tccggggcgt tatgtccgc attggtcttg accaactcta 18480
tcagagcttg gttgacggca atttcgatga tgcagcttg gcgcagggtc gatgcgacgc 18540
aatcgtccga tccggagccg ggactgtcgg gcgtacacaa atcgcccgca gaagcgcggc 18600
cgtctggacc gatggctgtg tagaagtact cgccgatagt ggaaaccgac gcccagcac 18660
tcgtccgagg gcaaaggaat agagtagatg ccgaccggg gatcgatcca cttaacgtta 18720
ctgaaatcat caaacagctt gacgaatctg gatataagat cgttggtgtc gatgtcagct 18780
ccggagttga gacaaatggg gttcaggatc tcgataagat acgttcattt gtccaagcag 18840
caaagagtgc cttctagtga tttaatagct ccatgtcaac aagaataaaa cgcgttttcg 18900
ggtttacctc ttccagatac agctcatctg caatgcatta atgcattgac tgcaacctag 18960
taacgccttn caggctccgg cgaagagaag aatagcttag cagagctatt ttcattttcg 19020
ggagacgaga tcaagcagat caacggtcgt caagagacct acgagactga ggaatccgct 19080
cttggtcca cgcgactata tatttgtctc taattgtact ttgacatgct cctcttctt 19140
actctgatag cttgactatg aaaattccgt caccagcncc tgggttcgca aagataattg 19200
catgtttctt ccttgaactc tcaagcctac aggacacaca ttcacgtag gtataaacct 19260

cgaaatcant tctactaag atggtataca atagtaacca tgcattggttg cctagtgaat 19320
gctccgtaac acccaatacg ccggccgaaa cttttttaca actctcctat gagtcgttta 19380
cccagaatgc acaggtacac ttgttttagag gtaatccttc tttctagcta gaagtcctcg 19440
tgtactgtgt aagcgccac tccacatctc cactcgacct gcaggcatgc a 19491

<210> 46
<211> 21300
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (3471)..(3471)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3679)..(3679)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3770)..(3770)
<223> n is a, c, g, or t

<400> 46
gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60
cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180
cttctgatcc agtcgatcat ccgctgaagg cgcttttcgaa tctggttaag atccacgtct 240

tcggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agctttctca 300

gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360

gtggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg 420

atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480

tttgatttg ctttccaggc tgagactcta gcttgagca tagagggtcc tttggcttcc 540

aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga 600

gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660

ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720

tccgatagct ctgcaaaggc cgttgcacaa tgcaagtcaa acgttgctag cagttccagg 780

tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840

caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg 900

cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960

tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca 1020

gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140

gttgaccggt gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200

ggagtgacct agaggggtcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260

aaatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcatcggaga atatggagct 1500

tcatcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560

cgaaatagca tgccattaac ctaggtagag aagtccaatt gcttccgata tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggg gctctgcact cgacctgctg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctg ccgcccggg tgtcggcggg gttgacaagg tcgttgctc 1920

agtccaacat ttgttgccat attttctgc tctcccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040

taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100

gttcgagctt tcccacttca tcgcagcttg actaacagct acccgcttg agcagacatc 2160

accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220

gtctccgacc tgatgcagct ctcgaggggc gaagaatctc gtgctttcag cttcgatgta 2280

ggagggcggtg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcgcccgcg ctcccgattc cggaagtgtc tgacattggg 2400

gaattcagcg agagcctgac ctattgcatc tcccgccgtg cacaggtgtc cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgaggaggc catggatgcg 2520

atcgctgagg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580

ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640

tggcaaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700

atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgacgcggga ttccggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820

ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880

atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgacagg atcgccgcgg 2940

ctcggggcgt atatgctccg cattggtctt gaccaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggg cgatgcgacg caatcgctcc atccggagcc 3060

gggactgtcg ggcgtacaca aatcgccccg agaagcgcgg ccgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cggcccagca ctctccgag ggcaaaggaa 3180

tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagtgt agacaaatgg 3300

tgttcaggat ctgataaga tacgttcatt tgtccaagca gcaaagagtg ctttctagt 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480

gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540

tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600

atatttgtct ctaattgtac ttgacatgc tcctcttctt tactctgata gcttgactat 3660

gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720

ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttctactaa 3780

gatggtatac aatagtaacc atgcatgggt gcctagttaa tgctccgtaa cacccaatac 3840

gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggtaca 3900

cttgttttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960

ctccacatct ccactcgacc tgcaggcatg caagcttgaa ttcgagctcg gtacccgggg 4020

atcttttcgac actgaaatac gtcgagcctg ctccgcttgg aagcggcgag gagcctcgtc 4080

ctgtcacaac taccaacatg gagtacgata agggccagtt cggccagctc attaagagcc 4140

agttcatggg cgttggcatg atggccgtca tgcattctgta cttcaagtac accaacgctc 4200

ttctgatcca gtcgatcatc cgctgaaggc gctttcgaat ctgggttaaga tccacgtctt 4260

cggaagcca gcgactgggtg acctccagcg tccctttaag gctgccaaca gctttctcag 4320

ccagggccag cccaagaccg acaaggcctc cctccagaac gccgagaaga actggagggg 4380

tggtgtcaag gaggagtaag ctcccttattg aagtcggagg acggagcggg gtcaagagga 4440

tattcttcga ctctgtatta tagataagat gatgaggaat tggaggtagc atagcttcat 4500

ttggatttgc tttccagget gagactctag cttggagcat agagggctct ttggctttca 4560

atattctcaa gtatctcgag tttgaactta ttccctgtga accttttatt caccaatgag 4620

cattggaatg aacatgaatc tgaggactgc aatcgccatg aggttttcga aatacatccg 4680

gatgtcgaag gcttggggca cctgcgttgg ttgaatttag aacgtggcac tattgatcat 4740

ccgatagctc tgcaaagggc gttgcacaat gcaagtcaaa cgttgctagc agttccaggt 4800

ggaatgttat gatgagcatt gtattaaatc aggagatata gcatgatctc tagttagctc 4860

accacaaaag tcagacggcg taaccaaag tcacacaaca caagctgtaa ggatttcggc 4920

acggctacgg aagacggaga agccaccttc agtggactcg agtaccattt aattctattt 4980

gtgtttgatc gagacctaat acagccccta caacgaccat caaagtcgta tagctaccag 5040

tgaggaagtg gactcaaatc gacttcagca acatctctg gataaacttt aagcctaaac 5100

tatacagaat aagataggtg gagagcttat accgagctcc caaatctgtc cagatcatgg 5160

ttgaccggtg cctggatctt cctatagaat catccttatt cgttgaccta gctgattctg 5220

gagtgacca gagggctcatg acttgagcct aaaatccgcc gcctccacca tttgtagaaa 5280

aatgtgacga actcgtgagc tctgtacagt gaccggtgac tctttctggc atgcggagag 5340
acggacggac gcagagagaa gggctgagta ataagccact ggccagacag ctctggcggc 5400
tctgaggtgc agtggatgat tattaatccg ggaccggccg cccctccgcc ccgaagtgga 5460
aaggctggtg tgcccctcgt tgaccaagaa tctattgcat catcggagaa tatggagctt 5520
catcgaatca ccggcagtaa gcgaaggaga atgtgaagcc aggggtgtat agccgtcggc 5580
gaaatagcat gccattaacc taggtacaga agtccaattg cttccgatct ggtaaaagat 5640
tcacgagata gtaccttctc cgaagtaggt agagcgagta cccggcgcgt aagctcccta 5700
attggcccat ccggcatctg tagggcgctc aaatatcgtg cctctcctgc tttgcccgtt 5760
gtatgaaacc ggaaaggccg ctcaggagct ggccagcggc gcagaccggg aacacaagct 5820
ggcagtcgac ccatccggtg ctctgcactc gacctgctga ggtccctcag tccctggtag 5880
gcagctttgc cccgtctgtc cgcccggtgt gtcggcggg ttgacaaggt cgttgctca 5940
gtccaacatt tgttgccata ttttcctgct ctccccacca gctgctcttt tcttttctct 6000
ttcttttccc atcttcagta tattcatctt cccatccaag aacctttatt tcccctaagt 6060
aagtactttg ctacatccat actccatcct tcccatccct tattcctttg aacctttcag 6120
ttcgagcttt cccacttcat cgcagcttga ctaacagcta ccccgcttga gcagacatca 6180
ccatgtcaat actcacttat ctggaatttc atctctacta tacactacct gtccttgcg 6240
cattgtgttg gctgctaaag ccgtttcact cacagcaaga caatctcaag tataaatttt 6300
taatgttgat ggccgcctct accgcatcga tttgggacaa ttatatcgtt tatcatcgcg 6360
cttggtggta ctgtcctact tgtgttggtg ctgtcattgg ctatgtacct ctagaagaat 6420
acatgttctt tatcatcatg actttaatga ctgtcgcgtt ctcaaacttt gttatgcgtt 6480
ggcacttgca tactttcttt attagacca acacttcttg gaagcaaaca ctattagtac 6540

gccttggtgcc tgtttcagct ttattggcaa tcacttatca tgcttggcac ttgacactgc 6600

caaataaacc ttcattttat ggttcatgca tcctttggta tgcttgtcct gtgttggcta 6660

ttctttggct ggggtgctggc gaatatatct tgcgtcgacc tgtggctgtc cttttgtcta 6720

ttgttatccc tagtgtatac ctatgttggg ctgatatcgt cgctattagt gctggcacat 6780

ggcatatttc tcttagaaca agcactggca aaatggtagt acccgattta cctgtagaag 6840

aatgcctgtt ttttactttg atcaacacag tcttggtttt tgctacctgt gctatagacc 6900

gcgctcaggc catcctccat gtgagcgcgc gtaatacgac tcactatagg gcgaattgga 6960

gctccaccgc ggtggcggcc gctctagaac tagtggatcc cccgggctgc aggaattcgg 7020

cacgagctac atttcacaag cccgtgagcg gtgcaagcgc tctgccccac atcggccac 7080

ctcctcatct ccacgggtca tttgctgcta ccacgatgt gtggaagctg cagtcaatca 7140

gcgtcaaggc ccgccgcgtt gaactagccc gcgacatcac gcggcccaaa gtctgcctgc 7200

atgctcagcg gtgctcgta gttcggctgc gagtggcagc accacagaca gaggaggcgc 7260

tgggaaccgt gcaggctgcc ggcgcgggcg atgagcacag cgccgatgta gcactccagc 7320

agcttgaccg ggctatcgca gagcgtcgtg cccggcgcaa acgggagcag ctgtcatacc 7380

aggctgccgc cattgcagca tcaattggcg tgtcaggcat tgccatcttc gccacctacc 7440

tgagatttgc catgcacatg accgtgggcg gcgcagtgcc atggggtgaa gtggctggca 7500

ctctcctctt ggtggttggg ggcgcgctcg gcatggagat gtatgcccg tatgcacaca 7560

aagccatctg gcatgagtcg cctctgggct ggctgctgca caagagccac cacacacctc 7620

gcactggacc ctttgaagcc aacgacttgt ttgcaatcat caatggactg cccgccatgc 7680

tcctgtgtac ctttggcttc tggctgccca acgtcctggg ggccggcctgc tttggagcgg 7740

ggctgggcat cacgctatac ggcattggcat atatgtttgt acacgatggc ctggtgcaca 7800

ggcgctttcc caccgggccc atcgctggcc tgcctacat gaagcgctg acagtggccc 7860

accagctaca ccacagcggc aagtacgggtg gcgcgccctg gggatatgtc ttgggtccac 7920

aggagctgca gcacattcca ggtgcggcgg aggaggtgga gcgactggtc ctggaactgg 7980

actgggtccaa gcgggctcag gccatcctcc atctgtacaa atcatctgtt caaaatcaaa 8040

accctaaaca agccatttcc cttttccagc atgtcaaaga gctagcatgg gccttctgtc 8100

ttcctgacca aatgctcaac aatgaattgt ttgatgatct tactatcagc tgggatattt 8160

tacgtaaagc ctcaaagtca ttctatactg catctgccgt tttccaagt tatgtacgtc 8220

aagacttggg tgttctctat gctttctgca gagctaccga tgacctgtgc gatgatgaat 8280

ccaaatctgt tcaagaaaga agagaccaat tagatcttac tcgacaattt gttcgtgatc 8340

tctttagcca aaagaccagt gcgcctattg tgattgattg ggaattgtat caaaaccaac 8400

ttcctgcttc ttgtatatca gcctttagag cctttactcg ccttcgccat gtccttgaag 8460

tagaccctgt agaagaacta ttagatgggt acaaatggga tcttgagcgt cgtcctatcc 8520

ttgatgaaca agacttggag gcatactctg cttgtgtggc cagtagtgtg ggtgaaatgt 8580

gcacacgtgt gattcttgct caagaccaa aggaaatga tgcttgata attgaccgtg 8640

cacgtgagat ggggctgggt ctacaatacg ttaacattgc tcgagacatt gtgactgata 8700

gcgagactct gggtcgatgt tatctgcctc aacaatggct tagaaaagaa gaaacagaac 8760

aaatacagca aggcaacgcc cgtagcctag gtgatcaaag actgttgggc ttgtctctga 8820

agcttgtagg aaaggcagac gctatcatgg tgagagctaa gaagggcatt gacaagttgc 8880

cggcaaaactg tcaaggcggg gtacgagctg cttgccaaat atatgctgca attggatctg 8940

tactcaagca gcagaagaca acatataccta caagagctca tctaaaagga agcgaacgtg 9000

ccaagattgc tctgttgagt gtatacaacc tctatcaatc tgaagacaag cctgtggctc 9060

tccgtcaagc tagaaagatt aagagttttt ttgttgatta gtgaattttt gttttattta 9120

tgtctgatag ttcaataaag agacaacaca tacaatataa aatcattgtc tttaaagtgt 9180

aatttagtag agtgtaaagc ctgcattttt tttgtacgca taaacaatga gttcaccccg 9240

cttctgggtt tttaataatt atgtcaaact agggaaaatt cttttttttc tcttcgttct 9300

ttttttggct tgttgtggag tcacaggctt gtcttcagat tgatagaggt tgtatacact 9360

caacagagca atcttggcac gttcgcttcc ttttagatga gctctttag gatatgttgt 9420

cttctgctgc ttgagtacag atccaattgc agcatatact tggcaagcag ctcgtacacc 9480

gccttgacag ttgcccggca acttgtcaat gcccttctta gctctcacca tgatagcgtc 9540

tgcctttcct acaagcttgg cgtaatcatg gtcatactg tttcctgtgt gaaattgtta 9600

tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag cctgggggtgc 9660

ctaattgagt agctaactca cattaattgc gttgcgctca ctgcccgtt tccagtcggg 9720

aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg 9780

tattgggcca aagacaaaag ggcgacattc aaccgattga gggagggaag gtaaattattg 9840

acggaaatta ttcatataag gtgaattatc accgtcacccg acttgagcca tttgggaatt 9900

agagccagca aatcaccag tagcaccatt accattagca aggccggaaa cgtcaccaat 9960

gaaaccatcg atagcagcac cgtaatcagt agcgacagaa tcaagtttgc ctttagcgtc 10020

agactgtagc gcgttttcat cggcattttc ggtcatagcc cccttattag cgtttgccat 10080

cttttcataa tcaaaatcac cggaaccaga gccaccaccg gaaccgcctc cctcagagcc 10140

gccaccctca gaaccgccac cctcagagcc accaccctca gagccgccac cagaaccacc 10200

accagagccg ccgccagcat tgacaggagg cccgatctag taacatagat gacaccgcgc 10260

gcgataattt atcctagttt gcgcgctata ttttgttttc tatcgcgat taaatgtata 10320

attgcgggac tctaatacata aaaacccatc tcataaataa cgtcatgcat tacatgttaa 10380

ttattacatg cttaacgtaa ttcaacagaa attatatgat aatcatcgca agaccggcaa 10440

caggattcaa tcttaagaaa ctttattgcc aaatgtttga acgatcgggg atcatccggg 10500

tctgtggcgg gaactccacg aaaatatccg aacgcagcaa gatatcgcg tgcatctcgg 10560

tcttgccctgg gcagtcgccg ccgacgccgt tgatgtggac gccggggccg atcatattgt 10620

cgctcaggat cgtggcggtg tgcttgtcgg ccgttgctgt cgtaatgata tcggcacctt 10680

cgaccgcctg ttccgcagag atcccgtggg cgaagaactc cagcatgaga tccccgcgct 10740

ggaggatcat ccagccggcg tcccggaaaa cgattccgaa gcccaacctt tcatagaagg 10800

cggcgggtgga atcgaaatct cgtgatggca ggttgggcgt cgcttggtcg gtcatttoga 10860

accccagagt cccgctcaga agaactcgtc aagaaggcga tagaaggcga tgcgctgcga 10920

atcgggagcg gcgataccgt aaagcacgag gaagcgggtca gccattcgc cgccaagctc 10980

ttcagcaata tcacgggtag ccaacgctat gtcctgatag cggcccgcca caccagccg 11040

gccacagtcg atgaatccag aaaagcggcc attttccacc atgatattcg gcaagcaggc 11100

atcgccatgg gtcacgacga gatcatcgcc gtcgggcatg cgcgccttga gcctggcgaa 11160

cagttcggct ggcgcgagcc cctgatgctc ttcgtccaga tcatcctgat cgacaagacc 11220

ggcttccatc cgagtacgtg ctcgctcgat gcgatgtttc gcttggtggt cgaatgggca 11280

ggtagccgga tcaagcgtat gcagccgccg cattgcatca gccatgatgg atactttctc 11340

ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgcca atagcagcca 11400

gtcccttccc gcttcagtga caacgtcgag cacagctgcg caaggaacgc ccgtcgtggc 11460

cagccacgat agccgcgctg cctcgtcctg cagttcattc agggcaccgg acaggtcgg 11520

cttgacaaaa agaaccgggc gccctgcgc tgacagccgg aacacggcgg catcagagca 11580

gccgattgtc tgttgtgccc agtcatagcc gaatagcctc tccaccaag cggccggaga 11640
acctgcgtgc aatccatctt gttcaatcat gcgaaacgat ccagatccgg tgcagattat 11700
ttggattgag agtgaatatg agactctaata tggataccga ggggaattta tggaacgtca 11760
gtggagcatt tttgacaaga aatatttgct agctgatagt gaccttaggc gacttttgaa 11820
cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa acccgcggt 11880
gagtggctcc ttcaacgttg cggttctgtc agttccaaac gtaaaacggc ttgtcccgg 11940
tcatcggcgg gggtcataac gtgactccct taattctccg ctcatgatca gattgtcggt 12000
tcccgccttc agtttaaact atcagtgttt gacaggatat attggcgggt aaacctaaga 12060
gaaaagagcg tttattagaa taatcggata tttaaaagg cgtgaaaagg tttatccgtt 12120
cgtccatttg tatgtgcatg ccaaccacag ggttccccag atctggcgcc ggccagcgag 12180
acgagcaaga ttggccgccc cccgaaacga tccgacagcg cgcccagcac aggtgcgcag 12240
gcaaattgca ccaacgcata cagcgccagc agaatgccat agtgggcggt gacgtcgttc 12300
gagtgaacca gatcgcgag gagggccggc agcaccggca taatcaggcc gatgccgaca 12360
gcgtcgagcg cgacagtgtc cagaattacg atcaggggta tgttgggttt cacgtctggc 12420
ctccggacca gcctccgtg gtccgattga acgcgaggat tctttatcac tgataagttg 12480
gtggacatat tatgtttatc agtgataaag tgtcaagcat gacaaagttg cagccgaata 12540
cagtgatccg tgccgccctg gacctgttga acgaggtcgg cgtagacggt ctgacgacac 12600
gcaaactggc ggaacggttg ggggttcagc agccggcgct ttactggcac ttcaggaaca 12660
agcgggcgct gctcgacgca ctggccgaag ccatgctggc ggagaatcat acgcattcgg 12720
tgccgagagc cgacgacgac tggcgctcat ttctgatcgg gaatgcccgc agcttcaggc 12780
aggcgctgct cgcctaccgc gatggcgcg gcacccatgc cggcacgcga cggggcgcac 12840

cgcagatgga aacggccgac ggcagcttc gcttctctg cgaggcgggt ttttcggccg 12900

gggacgccgt caatgcgctg atgacaatca gctacttcac tggtggggcc gtgcttgagg 12960

agcaggccgg cgacagcgat gccggcgagc gcggcgccac cgttgaacag gctccgctct 13020

cgccgctgtt gcggggccg atagacgcct tcgacgaagc cggtcgggac gcagcgttcg 13080

agcagggact cgcggtgatt gtcgatggat tggcgaaaag gaggctcggt gtcaggaacg 13140

ttgaaggacc gagaaagggt gacgattgat caggaccgct gccggagcgc aaccactca 13200

ctacagcaga gccatgtaga caacatcccc tcccccttc caccgctca gacgccgta 13260

gcagcccgct acgggctttt tcatgccctg ccctagcgtc caagcctcac ggccgcgctc 13320

ggcctctctg gcggccttct ggcgctcttc cgcttctcg ctactgact cgctgcgctc 13380

ggctgctcgg ctgcggcgag cggtatcagc tactcaaag gcgtaatac gggtatccac 13440

agaatcagg gataacgcag gaaagaacat gtgagcaaaa ggccagcaa aggccaggaa 13500

ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cccccctg acgagcatca 13560

caaaaatcga cgctcaagtc agaggtggcg aaaccgcaca ggactataaa gataccaggc 13620

gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 13680

cctgtccgcc tttctccctt cgggaagcgt ggcgcttttc cgctgcataa ccctgcttcg 13740

gggtcattat agcgattttt tcggtatata catccttttt cgcacgatat acaggatttt 13800

gccaaagggt tcgtgtagac tttccttgggt gtatccaacg gcgtcagccg ggcaggatag 13860

gtgaagtagg cccaccgcg agcgggtgtt ccttcttcac tgteccctat tcgcacctgg 13920

cggtgctcaa cgggaatcct gctctgcgag gctggccggc taccgccggc gtaacagatg 13980

agggcaagcg gatggctgat gaaaccaagc caaccaggaa gggcagccca cctatcaagg 14040

tgtaactgcct tccagacgaa cgaagagcga ttgaggaaaa ggccggcgcg gccggcatga 14100

gcctgtcggc ctacctgctg gccgtcggcc agggctacaa aatcacgggc gtcgtggact 14160
atgagcacgt ccgcgagctg gcccgcacatca atggcgacct gggccgcctg ggcggcctgc 14220
tgaaactctg gctcacccgac gaccgcgcga cggcgcggtt cggatgatgcc acgatcctcg 14280
ccctgctggc gaagatcgaa gagaagcagg acgagcttgg caaggatcatg atgggcgtgg 14340
tccgcccagag ggcagagcca tgactttttt agccgctaaa acggccgggg ggtgcgcgtg 14400
attgccaaagc acgtcccat gcgtccatc aagaagagcg acttcgcgga gctggtgaag 14460
tacatcacccg acgagcaagg caagaccgag cgcctttgcg acgtcacccg ggctggttgc 14520
cctcgccgct gggctggcgg ccgtctatgg ccctgcaaac gcgccagaaa cgcgctcgaa 14580
gccgtgtgcg agacaccgcg gccgcggcg ttgtggatac ctgcggaaa acttggccct 14640
cactgacaga tgagggggcg acgttgacac ttgaggggccc gactcaccg gcgcggcggtt 14700
gacagatgag gggcaggctc gatttcggcc ggcgacgtgg agctggccag cctcgcaaatt 14760
cggcgaaaac gcctgatttt acgcgagttt cccacagatg atgtggacaa gcctggggat 14820
aagtgcctg cggatttgac acttgagggg cgcgactact gacagatgag gggcgcgatc 14880
cttgacactt gagggggcaga gtgtgacag atgagggggc cacctattga catttgaggg 14940
gctgtccaca ggcagaaaat ccagcatttg caagggttcc cgcgcgtttt tcggccaccg 15000
ctaacctgtc ttttaacctg cttttaaac aatatttata aaccttggtt ttaaccaggg 15060
ctgcgccctg tgccgctgac cgcgcacgcc gaaggggggt gccccccctt ctgaaccct 15120
cccgcccgcc taacgcgggc ctcccatccc ccagggggt gcgcccctcg gccgcgaacg 15180
gcctcacccc aaaaatggca gcgtggcag tccttgccat tgccgggatc ggggcagtaa 15240
cgggatgggc gatcagcccg agcgcgacgc ccggaagcat tgacgtgccg cagggtgctg 15300
catcgacatt cagcgaccag gtgccgggca gtgagggcg cggcctgggt ggcggcctgc 15360

ccttcacttc ggccgtcggg gcattcacgg acttcatggc ggggcccggca atttttacct 15420
tgggcattct tggcatagt gtcgcgggtg ccgtgctcgt gttcgggggt gcgataaacc 15480
cagcgaacca tttgaggtga taggtaagat tataccgagg tatgaaaacg agaattggac 15540
ctttacagaa ttactctatg aagcgccata tttaaaaagc taccaagacg aagaggatga 15600
agaggatgag gaggcagatt gccttgaata tattgacaat actgataaga taatatatct 15660
tttatataga agatatcgcc gtatgtaagg atttcagggg gcaaggcata ggcagcgcgc 15720
ttatcaatat atctatagaa tgggcaaagc ataaaaactt gcatggacta atgcttgaaa 15780
cccaggacaa taaccttata gcttgtaaata tctatcataa ttgggtaatg actccaactt 15840
attgatagtg ttttatgttc agataatgcc cgatgacttt gtcatgcagc tccaccgatt 15900
ttgagaacga cagcgacttc cgtcccagcc gtgccagggt ctgcctcaga ttcaggttat 15960
gccgctcaat tcgctgcgta tatcgcttgc tgattacgtg cagctttccc ttcaggcggg 16020
attcatacag cggccagcca tccgtcatcc atatcaccac gtcaaagggt gacagcaggc 16080
tcataagacg ccccagcgtc gccatagtgc gttcaccgaa tacgtgcgca acaaccgtct 16140
tccggagact gtcatacgcg taaaacagcc agcgtggcg cgatttagcc ccgacatagc 16200
cccactgttc gtccatttcc ggcagacga tgacgtcact gcccggctgt atgcgcgagg 16260
ttaccgactg cggcctgagt tttttaagt acgtaaaatc gtgttgaggc caacgcccac 16320
aatgcgggct gttgcccggc atccaacgcc attcatggcc atatcaatga tttctgtgtg 16380
cgtaccgggt tgagaagcgg tgtaagtga ctgcagttgc catgttttac ggcagtgaga 16440
gcagagatag cgctgatgtc cggcgggtgct tttgccgtta cgcaccaccc cgtcagtagc 16500
tgaacaggag ggacagctga tagacacaga agccactgga gcacctcaaa aacaccatca 16560
tactactaaat cagtaagttg gcagcatcac ccataattgt ggtttcaaaa tcgggtccgt 16620

cgatactatg ttatacgcca actttgaaaa caactttgaa aaagctgttt tctgggtattt 16680

aaggttttag aatgcaagga acagtgaatt ggagttcgtc ttgttataat tagcttcttg 16740

gggtatcttt aaatactgta gaaaagagga aggaaataat aaatggctaa aatgagaata 16800

tcaccggaat tgaaaaaact gatcgaaaaa taccgctgcg taaaagatac ggaaggaatg 16860

tctcctgcta aggtatataa gctgggtgga gaaaatgaaa acctatattt aaaaatgacg 16920

gacagccggt ataaaggac cacctatgat gtggaacggg aaaaggacat gatgctatgg 16980

ctggaaggaa agctgcctgt tccaaaggtc ctgcactttg aacggcatga tggctggagc 17040

aatctgctca tgagtgaggc cgatggcgtc ctttgctcgg aagagtatga agatgaacaa 17100

agccctgaaa agattatcga gctgtatgcg gagtgcacat ggctctttca ctccatcgac 17160

atatcggatt gtccctatac gaatagctta gacagccgct tagccgaatt ggattactta 17220

ctgaataacg atctggccga tgtggattgc gaaaactggg aagaagacac tccatttaaa 17280

gatccgcgcg agctgtatga ttttttaaag acggaaaagc ccgaagagga acttgtcttt 17340

tcccacggcg acctgggaga cagcaacatc tttgtgaaag atggcaaagt aagtggcttt 17400

attgatcttg ggagaagcgg cagggcggac aagtggatg acattgcctt ctgctccgg 17460

tcgatcaggg aggatatcgg ggaagaacag tatgtcgagc tattttttga ctactgggg 17520

atcaagcctg attgggagaa aataaaatat tatattttac tggatgaatt gtttttagtac 17580

ctagatgtgg cgcaacgatg ccggcgacaa gcaggagcgc accgacttct tccgcatcaa 17640

gtgttttggc tctcaggccg agggccacgg caagtatttg ggcaaggggt cgctggtatt 17700

cgtgcagggc aagattcgga ataccaagta cgagaaggac ggccagacgg tctacgggac 17760

cgacttcatt gccgataagg tggattatct ggacaccaag gcaccaggcg ggtcaaatca 17820

ggaataaggg cacattgccc cggcgtgagt cggggcaatc ccgcaaggag ggtgaatgaa 17880

tccgacgttt gaccggaagg catacaggca agaactgata gacgcgggggt tttccgccga 17940

ggatgccgaa accatcgcaa gccgcaccgt catgcgtgcg ccccgcgaaa ccttcagtc 18000

cgtcggctcg atgggtccagc aagctacggc caagatcgag cgcgacagcg tgcaactggc 18060

tccccctgcc ctgcccgcg catcgggcgc cgtggagcgt tcgcgtcgtc tcgaacagga 18120

ggcggcaggt ttggcgaaat cgatgaccat cgacacgcga ggaactatga cgaccaagaa 18180

gcgaaaaacc gccggcgagg acctggcaaa acaggtcagc gaggccaagc aggccgcgtt 18240

gctgaaacac acgaagcagc agatcaagga aatgcagctt tccttgttcg atattgcgcc 18300

gtggccggac acgatgcgag cgatgcaaaa cgacacggcc cgctctgccc tggtcaccac 18360

gcgcaacaag aaaatcccgc gcgaggcgct gcaaaacaag gtcattttcc acgtcaacaa 18420

ggacgtgaag atcacctaca ccggcgctga gctgcggggc gacgatgacg aactgggtgtg 18480

gcagcaggtg ttggagtacg cgaagcgcac ccctatcggc gagccgatca ccttcacgtt 18540

ctacgagctt tgccaggacc tgggctggtc gatcaatggc cggattata cgaaggccga 18600

ggaatgcctg tcgcgcctac aggcgacggc gatgggcttc acgtccgacc gcgttgggca 18660

cctggaatcg gtgtcgctgc tgcaccgctt ccgcgtcctg gaccgtggca agaaaacgtc 18720

ccgttgccag gtctgatcg acgaggaaat cgtcgtgctg ttgctggcg accactacac 18780

gaaattcata tgggagaagt accgcaagct gtcgccgacg gcccgacgga tggtcgacta 18840

tttcagctcg caccgggagc cgtaccgct caagctggaa accttcgcc tcatgtgcgg 18900

atcggtattc acccgctga agaagtggcg cgagcaggtc ggcgaagcct gcgaagagtt 18960

gcgaggcagc ggcttgggtg aacacgcctg ggtcaatgat gacctgggtc attgcaaacg 19020

ctagggcctt gtggggtcag ttccggctgg gggttcagca gccagcgctt tactggcatt 19080

tcaggaacaa gcgggcactg ctgcacgcac ttgcttcgct cagtatcgct cgggacgcac 19140

ggcgcgctct acgaactgcc gataaacaga ggattaaaat tgacaattgt gattaaggct 19200

cagattcgac ggcttgagc ggccgacgtg caggatttcc gcgagatccg attgtcggcc 19260

ctgaagaaag ctccagagat gtccgggtcc gtttacgagc acgaggagaa aaagcccatg 19320

gaggcggtcg ctgaacggtt gcgagatgcc gtggcattcg gcgcctacat cgacggcgag 19380

atcattgggc tgtcgggtctt caaacaggag gacggcccca aggacgctca caaggcgcatt 19440

ctgtccggcg ttttcgtgga gcccgaaacag cgaggccgag gggtcgcggg tatgtctgtg 19500

cgggcgttgc cggcggggtt attgtctgtg atgatcgtcc gacagattcc aacgggaatc 19560

tggtggatgc gcatcttcat cctcggcgca cttaatattt cgctattctg gagcttgttg 19620

tttatttcgg totaccgcct gccggggcgg gtcgcggcga cggtaggcgc tgtgcagccg 19680

ctgatggtcg tgttcatctc tgccgctctg ctaggtagcc cgatacgatt gatggcggtc 19740

ctgggggcta tttgcggaac tgcgggcgtg gcgctgttgg tgttgacacc aaacgcagcg 19800

ctagatcctg tcggcgctgc agcgggcctg gcggggcggg tttccatggc gttcgggaacc 19860

gtgctgaccc gcaagtggca acctccctg cctctgctca cctttaccgc ctggcaactg 19920

gcggccggag gacttctgct cgttccagta gctttagtgt ttgatccgcc aatcccgatg 19980

cctacaggaa ccaatgttct cggcctggcg tggctcggcc tgatcggagc gggtttaacc 20040

taattccttt ggttccgggg gatctcgcga ctcgaaaccta cagttgtttc cttactgggc 20100

tttctcagcc ccagatctgg ggtcgtcag ccggggatgc atcaggccga cagtcggaac 20160

ttcgggtccc cgacctgtac cattcgggtga gcaatggata ggggagttga tatcgtcaac 20220

gttcacttct aaagaaatag cgccactcag ctctctcagc ggctttatcc agcgatttcc 20280

tattatgtcg gcatagttct caagatcgac agcctgtcac ggtaagcga gaaatgaata 20340

agaaggctga taattcggat ctctgcgagg gagatgatat ttgatcacag gcagcaacgc 20400

tctgtcatcg ttacaatcaa catgctaccc tccgcgagat catccgtggt tcaaaccg 20460
cagcttagtt gccgttcttc cgaatagcat cggtaacatg agcaaagtct gccgccttac 20520
aacggctctc ccgctgacgc cgtcccgga cgtatgggctg cctgtatcga gtggtgattt 20580
tgtgccgagc tgcgggtcgg ggagctgttg gctggctggg ggcaggatat attgtggtgt 20640
aaacaaattg acgcttagac aacttaataa cacattgcgg acgtttttaa tgtactgggg 20700
tggtttttct tttcaccagt gagacgggca acagctgatt gcccttcacc gcctggccct 20760
gagagagttg cagcaagcgg tccacgctgg tttgcccag caggcgaaaa tctgtttga 20820
tggtggttcc gaaatcggca aaatccctta taaatcaaaa gaatagcccg agatagggtt 20880
gagtgttgtt ccagtttgga acaagagtcc actattaaag aacgtggact ccaacgtcaa 20940
agggcgaaaa accgtctatc agggcgatgg ccactacgt gaaccatcac ccaaatcaag 21000
ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac cctaaaggga gccccgatt 21060
tagagcttga cggggaaagc cggcgaacgt ggcgagaaag gaagggaaga aagcgaaagg 21120
agcgggcgcc attcaggctg cgcaactgtt ggggaagggcg atcggtgagg gcctcttcgc 21180
tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag 21240
ggttttccca gtcaacgacgt tgtaaaacga cggccagtga attcgagctc ggtaccggg 21300

<210> 47

<211> 17756

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 47
ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgccg 120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240
cagcctcgca aatcggcgaa aacgcctgat tttacgcgag tttcccacag atgatgtgga 300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360
gagggggcgc atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480
ttttcggcca ccgctaacct gtottttaac ctgcttttaa accaatattt ataaccttg 540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600
cttctcgaac cctcccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc 660
tcggccgcga acggcctcac ccaaaaatg gcagcgctgg cagtccttgc cattgcggg 720
atcggggcag taacgggatg ggcgatcagc ccgagcgga cgcccgaag cattgacgtg 780
ccgcagggtg tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840

ggtaggggcc tgcccttcac ttcggccgctc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggctcggg gtgccgtgct cgtgttcggg 960

ggtagcdata acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactcaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac caggtcaaag 1560

ggtagacagca ggctcataag acgccccagc gtcgcatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatat gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttctgtccatt tccgcgcaga cgatgacgtc actgcccggc 1740

tgtagcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgtgat gtccggcggt gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggtttca 2100

aaatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc ttaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtgg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgaag 3360

gaggggtgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggatcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggg gtggcagcag gtgttggagt acgcaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc-tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggccccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgagagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctg 4500

tgcattgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcg cttacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcggggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctgggtgga tgcgcattct catcctcggc gcacttaata tttcgctatt 5100

ctggagcttg ttgtttatth cggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgt ctgctaggta gcccgatacg 5220

attgatggcg gtcctggggg ctatthtcgg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaagca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagt gcaacctccc gtgcctctgc tcacctttac 5400

cgcttgga ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttgggtccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcacaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttccact tctaaagaaa tagcgccact cagcttctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca togttacaat caacatgcta ccctccgcga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgtga cgccgtcccg gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120

tatatgttgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggctgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcgggtacc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaatac 7920

tgccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaatac cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgcttttggc cgggtgtatga aaccggaaaag gccgctcagg agctggccag cggcgcagac 8580

cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640

tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttgttgc catattttcc tgctctccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttccoct aagtaagtac tttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccatc tcatcgagc ttgactaaca gctacccgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctctgtcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgccgct gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgca ttgtgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggctgcc aacatcttct tctggaggcc 9660

gtgggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttggggcgag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcggggcgtag acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc a cactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncagggt cgggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgcgggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcggtgtact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt cattttgctt 10800

tgtaaatttc tggtaactgc caccaagaaa tatgaggata ttcgatgtgt tcctcgtggt 10860

agccaaaatg atagcacgtg ataaatgacc accaaatagg acggctaatt gtttgggcac 10920

aatgaggctg aacataaccc cctattgggt cactatgggg taaaaaagta ccaaaataga 10980

ataattgtaa tgaacttaaa agcgagggtg gcacccaaaa gtaagttaga ttatcacttg 11040

ggatatggag tatgtattta gcaaagttat aaataatagt caacgcaatt atttgccccc 11100

aactccagta acctttcata aaatgaaaat accaagcaaa gaaacttttg tgtttaccat 11160

tgtgaaaatc cgggtctatt gagcttgctg gattgtgggt gtgtaaccaa tgttttttca 11220

atagtttttg atatggtaaa agaccataaa gggatagggt caatgttcca atcaaatgat 11280

taatcttggg gttttgggga aatactacgc catgcatggc atcatgagat gtaataaata 11340

atcccgtata taaaaatgtt tgccatagta taacaggcaa taacatccaa aatttttagct 11400

ttgagatgtc aagggaagt aataaactca ggctaatac ccatgcgcta acaatgacaa 11460

tagcaatgaa aagcccctta aactgagatt tacttctcag tactggagtc agttttgctt 11520

gatgactgag tggttgttct aactggatca tttctaaaga gaaggtggaa caatgttagc 11580

ataattgtgc ttgagtgagg actttgaggg taggtacata cttgataaag ttaatgatta 11640

aagagaaaaa aaaagttttg gttcaaagca gaaattgttt tttaaactga ttggtgagaa 11700

aatttttttc tgtttccgca tcaccaaagc cacctcagga atggtcacaa attattggtc 11760

tgattggacc ataagcatac aaaaagttca ttgaagtata cttagtggct tattagactt 11820

ttatcgtttt ctaacgcgaa tcagcaatgt ttcttgtttg atttactgct tgcttttagat 11880

catttttgtc tgaaatatta tgcatttggt caaagcggcc tttgtttcct ttctttcatg 11940

cttaaacacg ttgtttattc catatattac tttgaatatg catcacgca aagcggaagt 12000

gcaaaaataac aaagaacctc tttgggttac acgatcaact gctattgtga aaaaaatttc 12060

tttttgaaaa tttttggaat aatatctctt gcaaaaaaga aattttgtat atttagtagc 12120

atcaagaaca aatgaaagaa gtgtgggata acaagaatac atcatcttta gacaaaagta 12180

cgagaaaaat ctaataagtt gttatagagg tctttgtttt ctttgtgttt atagacagtt 12240

atttagagtt tgaaaagtgt ctctaagtgt tcttttttta ttattattat ttcaaagtgt 12300

atgtaatata gctaaagcta tagatttgac attttttcta aatataaaat ttcagtcaac 12360

agaaataaat gacacgagtt ctttttctct ctctcaatcc tgttgatcat caatctttga 12420

tgtcgtttta aaacaaatga atggcattta gttccttagg tgtcactcac atcttgttga 12480

ccagaaaatc cttattcgcc ctcaaactct ctttattcct ttcatttgat ttgatgttta 12540

agtaatgcaa gcaaacaaaa aagaaacctt tcttgcaaag acaaaagaat tgttttcaga 12600

ggaaagcaac tcgttgatcat tttttaagga ttttagactta taatcgacac catagtttgt 12660

ccgttacatt ttttattgtc gttttctgat ttccttttaa tctttaagca aaatcaatat 12720

taacttatct tgtcttccaa taaaaaatgg ataccaataa caataaatcc ttcacaaaga 12780

aaaaaaaaaa aaactcgaaa aaagcttggc gtaatcatgg tcatagctgt ttcctgtgtg 12840

aaattgttat ccgctcacia ttccacacia catacgagcc ggaagcataa agtgtaaagc 12900

ctgggggtgcc taatgagtga gctaactcac attaatgtcg ttgcgctcac tgcccgtttt 12960

ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg 13020

cggtttgctg attgggccaa agacaaaagg gcgacattca accgattgag ggagggaagg 13080

taaatattga cggaaattat tcattaaagg tgaattatca ccgtcaccga cttgagccat 13140

ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa ggccggaaac 13200

gtcaccaatg aaaccatcga tagcagcacc gtaatcagta gcgacagaat caagtgtgcc 13260

tttagcgtca gactgtagcg cgttttcatc ggcattttcg gtcatagccc ccttattagc 13320

gtttgccatc ttttcataat caaaatcacc ggaaccagag ccaccaccgg aaccgcctcc 13380

ctcagagccg ccaccctcag aaccgccacc ctcagagcca ccaccctcag agccgccacc 13440

agaaccacca ccagagccgc cgccagcatt gacaggaggc ccgatctagt aacatagatg 13500

acaccgcgcg cgataattta tcctagtttg cgcgctatat tttgttttct atcgcgtatt 13560

aaatgtataa ttgcgggact ctaatcataa aaacccatct cataaataac gtcatgcatt 13620

acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata atcatcgcaa 13680

gaccggcaac aggattcaat cttaagaaac tttattgcc aatgtttgaa cgatcgggga 13740

tcatccgggt ctgtggcggg aactccacga aaatatccga acgcagcaag atatcgcggt 13800

gcatctcggt cttgcctggg cagtcgccgc cgacgcggt gatgtggacg ccgggcccga 13860

tcatatgtgc gtcaggatc gtggcggtgt gcttgctggc cgttgctgtc gtaatgatat 13920

cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc gaagaactcc agcatgagat 13980

ccccgcgctg gaggatcatc cagccggcgt cccggaaaac gattccgaag cccaaccttt 14040

catagaaggc ggcgggtggaa tcgaaatctc gtgatggcag gttgggcgtc gcttggtcgg 14100

tcatttcgaa ccccagagtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat 14160

gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag ccatttcgcc 14220

gccaaagtct tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac 14280

accagccgg ccacagtga tgaatccaga aaagcggcca tttccacca tgatattcgg 14340

caagcaggca tcgccatggg tcacgacgag atcatcgccg tcgggcatgc gcgccttgag 14400

cctggcgaac agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc 14460

gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc 14520

gaatgggcag gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga 14580

tactttctcg gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa 14640

tagcagccag tcccttcccc cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc 14700

cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc agttcattca gggcaccgga 14760
caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc 14820
atcagagcag ccgattgtct gttgtgcca gtcatagccg aatagcctct ccaccaagc 14880
ggcggagaaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgatc cagatccggt 14940
gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag gggaatttat 15000
ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg accttaggcg 15060
acttttgaac gcgcaataat ggtttctgac gtatgtgctt agctcattaa actccagaaa 15120
cccgcggctg agtggtcct tcaacgttgc ggttctgtca gttccaaacg taaaacggct 15180
tgtcccgct catcgccggg ggtcataacg tgactccctt aattctccgc tcatgatcag 15240
attgtcgttt cccgccttca gtttaacta tcagtgttg acaggatata ttggcgggta 15300
aacctaagag aaaagagcgt ttattagaat aatcgatat taaaagggc gtgaaaagg 15360
ttatccgttc gtccatttgt atgtgcatgc caaccacagg gttccccaga tctggcgccg 15420
gccagcgaga cgagcaagat tggccgccgc ccgaaacgat ccgacagcgc gcccagcaca 15480
ggtgcgcagg caaattgcac caacgcatac agcgccagca gaatgccata gtgggcgggtg 15540
acgtcgttcg agtgaaccag atcggcagg aggcccgga gcaccggcat aatcaggccg 15600
atgccgacag cgtcgagcgc gacagtgtc agaattacga tcaggggtat gttgggtttc 15660
acgtctggcc tccggaccag cctccgctgg tccgattgaa cgcgcggtt ctttatcact 15720
gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg acaaagttgc 15780
agccgaatac agtgatccgt gccgccctgg acctgttgaa cgaggtcggc gtagacggtc 15840
tgacgacacg caaactggcg gaacggttg gggttcagca gccggcgctt tactggcact 15900
tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc catgctggcg gagaatcata 15960

cgcattcggg gccgagagcc gacgacgact ggcgctcatt tctgatcggg aatgcccgc 16020
gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg catccatgcc ggcacgcgac 16080
cgggcgccacc gcagatggaa acggccgacg cgcagcttcg cttcctctgc gaggcgggtt 16140
tttcggcccg ggacgccgtc aatgcgctga tgacaatcag ctacttcact gttggggccg 16200
tgcttgagga gcaggccggc gacagcgatg ccggcgagcg cggcggcacc gttgaacagg 16260
ctccgctctc gccgctgttg cgggcccgcga tagacgcctt cgacgaagcc ggtccggacg 16320
cagcgttcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg aggctcgttg 16380
tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc aggaccgctg ccggagcgca 16440
accactcac tacagcagag ccatgtagac aacatcccct cccctttcc accgcgtcag 16500
acgccgtag cagcccgtc cgggcttttt catgccctgc cctagcgtcc aagcctcacg 16560
gccgcgctcg gcctctctgg cggccttctg gcgctcttcc gcttcctcgc tctactgactc 16620
gctgcgctcg gtcgttcggc tcggcgagc ggtatcagct cactcaaagg cggtaatagc 16680
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 16740
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 16800
cgagcatcac aaaaatcgac gctcaagtca gagtgggcga aaccgacag gactataaag 16860
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 16920
taccggatac ctgtccgctt ttctcccttc gggaagcgtg gcgcttttcc gctgcataac 16980
cctgcttcgg ggtcattata gcgatttttt cggtatatcc atcctttttc gcacgatata 17040
caggattttg ccaaaggggt cgtgtagact ttccttggtg tatccaacgg cgtcagccgg 17100
gcaggatagg tgaagtaggc ccaccgcga gcgggtgttc cttcttcact gtcccttatt 17160
cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg ctggccggct accgccggcg 17220

taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag ggcagcccac 17280
ctatcaaggt gtactgcctt ccagacgaac gaagagcgat tgaggaaaag gggcgggcgg 17340
ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca gggctacaaa atcacgggcg 17400
tcgtggacta tgagcacgtc cgcgagctgg ccgcatcaa tggcgacctg ggccgcctgg 17460
gcggcctgct gaaactctgg ctaccgacg acccgcgcac ggcgcgggtc ggtgatgcca 17520
cgatcctcgc cctgctggcg aagatcgaag agaagcagga cgagcttggc aaggatcatga 17580
tgggcgtggt ccgcccgagg gcagagccat gactttttta gccgctaaaa cggccggggg 17640
gtgcgcgtga ttccaagca cgtcccatg cgctccatca agaagagcga cttcgcggag 17700
ctggtgaagt acatcaccga cgagcaaggc aagaccgagc gcctttgcga cgctca 17756

<210> 48

<211> 17118

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 48

cgggctggt tgcctcgcc gctgggctgg cggcgtcta tggcctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gggccgccg gcgttggtga tacctgcgg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttcccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcgttatt gacacttgag gggcgcgact actgacagat	360
gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgaagggt ttccgccgt	480
tttcggcca cgcctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctgaac cctcccgcc cgctaacgcg ggctcccat cccccaggg gctgcgcccc	660
tggccgcga acggcctcac ccaaaaatg gcagcgctgg cagtccctgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgga cgccgggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg	840
ggtggcggcc tgccttcac ttggccgtc ggggcattca cggacttcat ggcggggccg	900
gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag	1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata	1140
agataatata tcttttatat agaagatac gccgtatgta aggatttcag ggggcaaggc	1200

ataggcagcg cgcttatcaa tataatctata gaatgggcaa agcataaaaa cttgcatgga 1260
ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttcggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgctccatt tccgcgcaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcggcctg agtttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggg gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggtttca 2100
aaatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaagggtt tagaatgcaa ggaacagtga attggagtgc gtcttggtat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcttgcaact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc tcaaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccgcgctg agtcggggca atcccgaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaacottcca gtccgtcggc tcgatggtcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccc cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgcg 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcattgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcgggccgc gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggctg tgccggcggtg tttattgctc gtgatgatcg tccgacagat 5040
tccaaacggga atctggtgga tgcgcattct catcctcggc gcacttaata tttcgctatt 5100
ctggagcttg ttgtttatct cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtcttggggg ctatttgctg aactgcgggc gtggcgctgt tgggtgttgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgcttgcaa ctggcgggcg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttcgg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcattcaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacgggtaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctcgc agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgca gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgct tacaacggct ctcccgtga cggcgctccg gactgatggg ctgcctgtat 6060
cgagtgggga tttgtgccc agctgccggt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgcttcc gaatctgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccgatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcggttga caatgcaagt caaacggtgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtga ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaadc 7920
tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcgttga 7980
cctagctgat tctggagtga ccagagggt catgacttga gcctaaaadc cgccgctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactcttcc 8100
tggtcatgagg agagacggac ggacgcagag agaagggtg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgtatga aaccggaaag gccgctcagg agctggccag cggcgagac 8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggctcc 8640
tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggtcgttgc gtcagtcctc catttgttgc catattttcc tgctctcccc accagctgct 8760

ctttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttccctt aagtaagtac ttgtctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctcgtgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgtgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctccg tcgcgcaggg 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgaatgc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcggt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcggaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctcgg 10620

taacacccaa tacgcgggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggt acaattgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800

tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860

atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920

aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt tttttttat 10980

acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040

atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100

tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat tttactaatt 11160

cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220

ttgattacac tctaaatctc gatgaccgga ccaaaaagag cgaacaaaga aataatcttg 11280

tgcattegaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340

tgttatataa acaaaaagaa attgtacaaa cttctctttc ttctcttttt attttatctc 11400

tatgatccag ttagaacaac cactcagtca tcaagcaaaa ctgactccag tactgagaag 11460

taaactctcag ttttaagggc ttttcattgc tattgtcatt gttagcgcac gggtcattag 11520

cctgagttta ttactttccc ttgacatctc aaagctaaaa ttttgatgt tattgcctgt 11580

tatactatgg caaacatttt tatatacggg attatttatt acatctcatg atgccatgca 11640

tggcgtagta tttcccaaaa acaccaagat taatcatttg attggaacat tgaccctatc 11700

cctttatggc cttttaccat atcaaaaact attgaaaaaa cattgggtac accaccacaa 11760

tccagcaagc tcaatagacc cggattttca caatggtaaa caccaaagtt tctttgcttg 11820

gtattttcat tttatgaaag gttactggag ttgggggcaa ataattgcgt tgactattat 11880

ttataacttt gctaaatata tactccatat cccaagtcat aatctaactt acttttgggt 11940

gtaccctcg cttttaagtt cattacaatt attctatttt ggtacttttt taccocatag 12000

tgaaccaata ggggggtatg ttcagcctca ttgtgcccaa acaattagcc gtcctatttg 12060

gtggtcattt atcacgtgct atcatttttg ctaccacgag gaacatcacg aatatcctca 12120

tatttcttg tgccagttac cagaaattta caaagcaaaa tagaagcttg gcgtaatcat 12180

ggtcatagct gtttcctgtg tgaaattggt atccgctcac aattccacac aacatacgag 12240

ccggaagcat aaagtgtaaa gcctgggggtg cctaattgag gagctaactc acattaattg 12300

cgttgcgctc actgcccgtc ttccagtcg gaaacctgtc gtgccagctg cattaatgaa 12360

tcggccaacg cgcggggaga ggcgggttgc gtattgggcc aaagacaaaa gggcgacatt 12420

caaccgattg agggagggaa ggtaaatatt gacggaaatt attcattaaa ggtgaattat 12480

caccgtcacc gacttgagcc atttgggaat tagagccagc aaaatcacca gtagcaccat 12540

taccattagc aaggccggaa acgtcaccaa tgaaaccatc gatagcagca ccgtaatcag 12600
tagcgacaga atcaagtttg ccttttagcgt cagactgtag cgcgttttca tcggcatttt 12660
cggtcatagc ccccttatta gcggttgcca tcttttcata atcaaaatca ccggaaccag 12720
agccaccacc ggaaccgcct ccctcagagc cgccaccctc agaaccgcca ccctcagagc 12780
caccaccctc agagccgcca ccagaaccac caccagagcc gccgccagca ttgacaggag 12840
gcccgatcta gtaacataga tgacaccgcg cgcgataatt tctcctagtt tgcgcgctat 12900
attttgtttt ctatcgcgta ttaaattgtat aattgcgggga ctctaatacat aaaaacccat 12960
ctcataaata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13020
aattatatga taatcatcgc aagaccggca acaggattca atcttaagaa actttattgc 13080
caaatgtttg aacgatcggg gatcatccgg gtctgtggcg ggaactccac gaaaatatcc 13140
gaacgcagca agatatcgcg gtgcatctcg gtcttgcccg ggcagtcgcc gccgacgccg 13200
ttgatgtgga cgccggggccc gatcatattg tcgctcagga tcgtggcggtt gtgcttgctg 13260
gccgttgctg tcgtaatgat atcggcacct tcgaccgect gttccgcaga gatcccgtag 13320
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccggaag 13380
acgattccga agcccaacct ttcatagaag gcggcggtgg aatcgaaatc tcgtgatggc 13440
aggttgggcg tcgcttggtc ggtcatttcg aacccagag tcccgctcag aagaactcgt 13500
caagaaggcg atagaaggcg atgcgctcgc aatcgggagc ggcgataccg taaagcacga 13560
ggaagcggtc agcccatcgc ccgccaagct cttcagcaat atcacgggta gccaacgcta 13620
tgtcctgata gcggtccgcc acaccagcc ggccacagtc gatgaatcca gaaaagcggc 13680
cattttccac catgatattc ggcaagcagg catcgccatg ggacacgacg agatcatcgc 13740
cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc ccctgatgct 13800

cttcgtccag atcatcctga tcgacaagac cggcttccat ccgagtacgt gctcgtcga 13860
tgcatgttt cgcttggtgg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc 13920
gcattgcac agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat 13980
cctgccccgg cacttcgccc aatagcagcc agtcccttcc cgcttcagtg acaacgtcga 14040
gcacagctgc gcaaggaacg cccgtcgtgg ccagccacga tagccgcgct gcctcgtcct 14100
gcagttcatt cagggcaccg gacaggtcgg tcttgacaaa aagaaccggg cggccctgcg 14160
ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttggtgcc cagtcatagc 14220
cgaatagcct ctccacccaa gcggccggag aacctgcgtg caatccatct tgttcaatca 14280
tgcgaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14340
ttggataccg aggggaattt atggaacgtc agtggagcat ttttgacaag aaatatttgc 14400
tagctgatag tgaccttagg cgacttttga acgcgcaata atggtttctg acgtatgtgc 14460
ttagctcatt aaactccaga aaccgcggc tgagtggctc cttcaacgtt gcggttctgt 14520
cagttccaaa cgtaaaacgg cttgtccgc gtcacggcg ggggtcataa cgtgactccc 14580
ttaattctcc gctcatgac agattgtcgt ttcccgctt cagtttaaac tatcagtgtt 14640
tgacaggata tattggcggg taaacctaag agaaaagagc gtttattaga ataacggat 14700
atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gcccaaccaca 14760
gggttcccca gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg 14820
atccgacagc gcgccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag 14880
cagaatgcca tagtgggagg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg 14940
cagcaccggc ataatcaggc cgatgcgac agcgtcgagc gcgacagtgc tcagaattac 15000
gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15060

aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa 15120
gtgtcaagca tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg 15180
aacgaggtcg gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt gggggttcag 15240
cagccggcgc tttactggca cttcaggaac aagcggggcg tgctcgacgc actggccgaa 15300
gccatgctgg cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca 15360
tttctgatcg ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg 15420
cgcatccatg ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt 15480
cgcttctctc gcgagggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc 15540
agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag 15600
cgcgggcgga ccgttgaaca ggctccgctc tcgccgctgt tgccggccgc gatagacgcc 15660
ttcgacgaag ccggtccgga cgcagcgctc gagcagggac tcgcggtgat tgtcgatgga 15720
ttggcgaaaa ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga 15780
tcaggaccgc tgccggagcg caaccactc actacagcag agccatgtag acaacatccc 15840
ctcccccttt ccaccgcgct agacgcccg agcagccgcg tacgggcttt ttcatgccct 15900
gccctagcgt ccaagcctca cggccgcgct cggcctctct ggccggcctc tggcgctctt 15960
ccgcttctc gctcaactgac tcgctgcgct cggctcgctc gctgcggcga gcggtatcag 16020
ctcaactcaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 16080
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcttt 16140
tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc 16200
gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct 16260
ctcctgttcc gaccctgccg cttaccgat acctgtccgc ctttctcct tcgggaagcg 16320

tggcgctttt ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat 16380
ccatcctttt tcgcacgata tacaggattt tgccaaaggg ttctgttaga ctttccttgg 16440
tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt 16500
tcctttctca ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga 16560
ggctggccgg ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag 16620
ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16680
attgaggaaa aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc 16740
cagggctaca aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcac 16800
aatggcgacc tgggcgcctt gggcggcctg ctgaaactct ggctcaccga cgacccgcgc 16860
acggcgcggt tcggtgatgc cagatcctc gcctgctgg cgaagatcga agagaagcag 16920
gacgagcttg gcaaggatcat gatgggcgtg gtccgccga gggcagagcc atgacttttt 16980
tagccgctaa aacggccggg gggcgcgctg gattgccaag cacgtcccca tgcgctccat 17040
caagaagagc gacttcggg agctggtgaa gtacatcacc gacgagcaag gcaagaccga 17100
gcgcctttgc gacgctca 17118

<210> 49

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 49

gatctttcga cactgaaata cgtcgagcct gtcgcgcttg gaagcggcga ggagcctcgt	60
cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc	120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct	180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctgggtaag atccacgtct	240
tcggaagcc agcgactggt gacctccagc gtocctttaa ggctgccaac agctttctca	300
gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg	360
gtggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttggaattg ctttccaggc tgagactcta gcttgaggca tagagggtcc tttggcttcc	540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga	600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca	720
tccgatagct ctgcaaaggc cgctgcacaa tgcaagtcaa acgttgctag cagttccagg	780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840
caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg	900

cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960

tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca 1020

gtgaggaagt ggactcaaat cgacttcagc aacatctoct ggataaactt taagcctaaa 1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140

gttgaccggt gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200

ggagtgacct agaggggtcat gacttgagcc taaaatccgc cgctccacc atttgtagaa 1260

aatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc ccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcacgggaga atatggagct 1500

tcacgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagcgcgcg 1560

cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgatc tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccgg 1740

tgtatgaaac cggaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcaact cgacctgctg aggtccctca gtcctggta 1860

ggcagctttg ccccgctctgt ccgcccgggtg tgcgggcggg gttgacaagg tcgttgcgtc 1920

agtccaacat ttgttgccat attttctctc tctcccacc agctgctctt ttcttttctc 1980

ttctttttcc catcttcagt atattcatct tcccatcaa gaacctttat tccccctaag 2040

taagtacttt gctacatcca tactccatcc tcccatccc ttattccttt gaacctttca 2100

gttcgagctt tccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160

accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaataa gtctgacagc 2220

gtctccgacc tgatgcagct ctccggagggc gaagaatctc gtgctttcag ctccgatgta 2280

ggagggcgctg gatatgtcct gcgggtaaat agctgcgccg atgggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgtc tgacattggg 2400

gaattcagcg agagcctgac ctattgcacg tcccgccgtg cacaggggtg cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgaggaggc catggatgcg 2520

atcgctgcgg ccgatccttag ccagacgagc ggggttcggcc cattcggacc gcaaggaatc 2580

ggccaataca ctacatggcg tgatttcata tgccgcgattg ctgatcccca tgtgtatcac 2640

tggcaaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700

atgctttggg ccgaggactg ccccgaaagc cggcacctcg tgcacgcgga ttccggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcgggtc ttgactggag cgaggcgatg 2820

ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880

atggagcagc agacgcgcta ctccgagcgg aggcattccg agcttgacgg atcgccgcgg 2940

ctccgggctg atatgctccg cattgggtctt gaccaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgctcc atccggagcc 3060

gggactgtcg ggcgtacaca aatcgccgcg agaagcgcgg ccgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cgcgccagca ctggtccgag ggcaaaggaa 3180

tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300

tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagt ccttctagt 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480

gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540

tcaacggtcg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600

atatattgtct ctaattgtac ttgacatgc tctcttctt tactctgata gcttgactat 3660

gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720

ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttcctactaa 3780

gatggtatac aatagtaacc atgcatggtt gcctagttaa tgctccgtaa cacccaatac 3840

gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggatca 3900

cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960

ctccacatct cactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020

agaaagttaa aagaaattcg gaagcatggc acattcttct ttttataaat acatgcctga 4080

ctttcttttt ccatcgatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140

ggagtttgaa attttgcct ccaggagcaa aaaaaagttt ttttttatac atgtttgtac 4200

acaagaatag ttaccaattt gctttggtct tacgtgctgc aagtttatat cgttttcaat 4260

ttctttgtct ttacattttc ttgtccttt atctttctc atttagtctt tgggagaatt 4320

aggaaaagg agcggaagg taagaaatgc ttgcgtatct tactaattcg gcaaactcc 4380

aatttgcaa acagcagcct gtgcaacgct ctgagatga cagtatctt gattacactc 4440

taaatctcga tgacctgacc aaaaagagcg aacaaagaaa taatcttgtg cattcgaata 4500

tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560

aaaaagaaat tgtacaaact ttcttttctt ctctttttat tttatctcta tgctgtcgaa 4620

gctgcagtca atcagcgtca aggcccgccg cggtgaacta gcccgcgaca tcacgcggcc 4680

caaagtctgc ctgcatgctc agcgggtgctc gttagttcgg ctgcgagtgg cagcaccaca 4740

gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg ggcgatgagc acagcgccga 4800

tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccggc gcaaacggga 4860

gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920

cttcgccacc tacctgagat ttgccatgca catgaccgtg ggccggcgag tgccatgggg 4980

tgaagtggct ggcaactctcc tcttggtggt tgggtggcgcg ctgggcatgg agatgtatgc 5040

ccgctatgca cacaaagcca tctggcatga gtgcctctg ggctggctgc tgcacaagag 5100

ccaccacaca cctcgcactg gaccctttga agccaacgac ttgtttgcaa tcatcaatgg 5160

actgcccgcc atgctcctgt gtacctttgg cttctggctg cccaacgtcc tgggggcggc 5220

ctgctttgga gcggggctgg gcatcacgct atacggcatg gcatatatgt ttgtacacga 5280

tggcctggtg cacaggcgct ttcccaccgg gcccatcgct ggcttgccct acatgaagcg 5340

cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgcg cctgggggtat 5400

gttcttgggt ccacaggagc tgcagcacat tccaggtgcg gcggaggagg tggagcgact 5460

ggtcctggaa ctggactggt ccaagcgggc gattgtgact gatagcgaga ctctgggtcg 5520

atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaaatac agcaaggcaa 5580

cgcccgtagc ctaggtgatc aaagactggt gggcttgtct ctgaagcttg taggaaaggc 5640

agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700

cggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760

gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaga ttgctctggt 5820

gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880

gattaagagt ttttttggtg attagtgaat ttttgtttta tttatgtctg atagttcaat 5940

aaagagacaa cacatacaat ataaaatcat tgtctttaa tgtaattta gtagagtgt 6000

aagcctgcat tttttttgta cgcataaaca atgaattcac cccgcttctg gtttttaa 6060

aattatgtca aactaggga aattcttttt tttctcttcg ttcttttttt ggcttggtgt 6120

ggagtacacag gcttgtcttc agattgatag aggttgata cactcaacag agcaatcttg 6180

gcacgttcgc ttccttttag atgagctctt gtaggatatg ttgtcttctg ctgcttgagt 6240

acagatccaa ttgcagcata tacttgga gcagctcgta caccgccttg acagtttgcc 6300

ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tcctacaagc 6360

ttcagagaca agcccaacag tctttgatca cctaggctac gggcgttgcc ttgctgtatt 6420

tgttctgttt ctctttttct aagccattgt tgaggcagat aacatcgacc caacatcctc 6480

gagccatact acagcataaa aggatacgtt ttctttaaca gaaatttacc cttttgttat 6540

cagcacatac aaaaaaaaaag aaatttaaga tgagtaggac ttccattctc tcaaaaattt 6600

tattcaatcc ataatgaat tatttttggc caaaaagaa agattatgcc tgattttctc 6660

tatttttttt ttttttacia ctccaccaat actttctagc ccagcttggc gtaatcatgg 6720

tcatagctgt ttcctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc 6780

ggaagcataa agtgtaaagc ctgggggtgc taatgagtga gctaactcac attaattgag 6840

ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 6900

ggccaacgcg cggggagagg cggtttgcgt attgggcaa agacaaaagg gcgacattca 6960

accgattgag ggagggaagg taaatattga cggaaattat tcattaaagg tgaattatca 7020

ccgtcaccca cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080

ccattagcaa ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta 7140

gcgacagaat caagtttgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg 7200

gtcatagccc ccttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag 7260
ccaccaccgg aaccgcctcc ctgagagccg ccaccctcag aaccgccacc ctgagagcca 7320
ccaccctcag agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggagggc 7380
ccgatctagt aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat 7440
tttgttttct atcgcgatt aaatgtataa ttgcgggact ctaatcataa aaacccatct 7500
cataaataac gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560
ttatatgata atcatcgcaa gaccggcaac aggattcaat cttagaaac ttattgcca 7620
aatgtttgaa cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga 7680
acgcagcaag atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt 7740
gatgtggacg ccggggccga tcatattgtc gctcaggatc gtggcgttgt gcttgtcggc 7800
cgttgctgtc gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc 7860
gaagaactcc agcatgagat cccgcgctg gaggatcatc cagccggcgt cccgaaaac 7920
gattccgaag cccaacctt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag 7980
gttgggcgtc gcttggtcgg tcatctcgaa cccagagtc ccgctcagaa gaactcgtca 8040
agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100
aagcggtcag cccattcgcc gccaaactct tcagcaatat cacgggtagc caacgctatg 8160
tcctgatagc ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca 8220
ttttccacca tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg 8280
tcgggcatgc gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct 8340
tcgtccagat catcctgatc gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg 8400
cgatgtttcg cttggtggtc gaatgggcag gtagccggat caagcgtatg cagccgccgc 8460

attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520

tgccccggca cttcgcccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc 8580

acagctgcg c aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc 8640

agttcattca gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct 8700

gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgcca gtcatagccg 8760

aatagcctct ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820

cgaaacgatc cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880

ggataccgag gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940

gctgatagt accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000

agctcattaa actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca 9060

gttccaaacg taaaacggct tgtcccgcgt catcgccggg ggtcataacg tgactccctt 9120

aattctccgc tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg 9180

acaggatata ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcggatat 9240

ttaaaagggc gtgaaaaggt ttatccgttc gtccatttgt atgtgcatgc caaccacagg 9300

gttccccaga tctggcgccg gccagcgaga cgagcaagat tggccgccc cggaaacgat 9360

ccgacagcgc gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca 9420

gaatgccata gtgggcggtg acgtcgttcg agtgaaccag atcgcgagg aggcccgga 9480

gcaccggcat aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga 9540

tcaggggtat gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa 9600

cgcgcggtt ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt 9660

gtcaagcatg acaaagttgc agccgaatac agtgatecgt gccgccctgg acctgttgaa 9720

cgaggtcggc gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca 9780

gccggcgctt tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc 9840

catgctggcg gagaatcata cgcattcggc gccgagagcc gacgacgact ggcgctcatt 9900

tctgatcggg aatgcccga gcttcaggca ggcgctgctc gcctaccgag atggcgcgcg 9960

catccatgcc ggcacgcgac cgggcgaccc gcagatggaa acggccgacg cgcagcttcg 10020

cttcctctgc gaggcgggtt tttcggccgg ggacgccgtc aatgcgctga tgacaatcag 10080

ctacttcaact gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg 10140

cggcggcacc gttgaacagg ctccgctctc gccgctgttg cgggccgga tagacgcctt 10200

cgacgaagcc ggtccggacg cagcgcttca gcagggactc gcggtgattg tcgatggatt 10260

ggcgaaaagg aggctcgttg tcaggaacgt tgaaggaccg agaaagggcg acgattgatc 10320

aggaccgctg ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct 10380

ccccctttcc accgcgtcag acgccgtag cagcccgtca cgggcttttt catgccctgc 10440

cctagcgctc aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc 10500

gcttctctgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 10560

caactaaagg cggtaatagc gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620

tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 10680

cataggtctc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 10740

aaccgcagag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 10800

cctgttccga ccctgccgct taccggatac ctgtccgctt ttctcccttc gggaagcgctg 10860

gcgcttttcc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc 10920

atcctttttc gcacgatata caggattttg ccaaaggggt cgtgtagact ttccttggtg 10980

tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccaccgcga gcgggtgttc 11040

cttcttcact gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg 11100

ctggccgggt accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160

aaccaggaag ggcagccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat 11220

tgaggaaaag gcggcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 11280

gggctacaaa atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 11340

tggcgacctg ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac 11400

ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 11460

cgagcttggc aaggatcatga tgggcgtggt ccgcccaggg gcagagccat gactttttta 11520

gccgctaaaa cgcccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca 11580

agaagagcga cttcgcgagg ctggtgaagt acatcaccca cgagcaaggc aagaccgagc 11640

gcctttgcca cgctcacccg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc 11700

cctgcaaacg cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt 11760

tgtggatacc tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact 11820

tgagggggcg actcacccgg cgcggcgctg acagatgagg ggcaggctcg atttcggccg 11880

gcgacgtgga gctggccagc ctcgcaaatc ggcgaaaacg cctgatttta cgcgagtttc 11940

ccacagatga tgtggacaag cctgggggata agtgccctgc ggtattgaca cttgaggggc 12000

gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 12060

tgagggggcg acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc 12120

aagggtttcc gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca 12180

atatttataa accttgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcagccg 12240

aaggggggtg ccccccttc tcgaaccctc ccggcccgtt aacgcgggcc tcccatcccc 12300
ccagggggtg cgcctctcg cgcgaacgg cctcacccca aaaatggcag cgctggcagt 12360
ccttgccatt gccgggatcg gggcagtaac gggatggcg atcagcccga gcgcgacgcc 12420
cgggaagcatt gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag 12480
tgagggcgcc gccctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga 12540
cttcatggcg gggccggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc 12600
cgtgctcgtg ttccgggggtg cgataaaccc agcgaacccat ttgaggtgat aggtgaagatt 12660
ataccgaggt atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat 12720
ttaaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat 12780
attgacaata ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga 12840
tttcaggggg caaggcatag gcagcgcgt tatcaatata tctatagaat gggcaaagca 12900
taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt 12960
ctatcataat tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgccc 13020
gatgactttg tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg 13080
tgccaggtgc tgctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgc 13140
gattacgtgc agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca 13200
tatcaccacg tcaaagggtg acagcaggct cataagacgc ccagcgctcg ccatagtgcg 13260
ttcaccgaat acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca 13320
gcgctggcgc gatttagccc cgacatagcc ccaactgttcg tccatctccg cgcagacgat 13380
gacgtcactg cccggctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga 13440
cgtaaaatcg tgttgaggcc aacgcccata atgcgggctg ttgcccggca tccaacgcca 13500

ttcatggcca tatcaatgat tttctggtgc gtaccgggtt gagaagcggg gtaagtgaac 13560
tgcaagttgcc atgttttacg gcagtgaagag cagagatagc gctgatgtcc ggcgggtgctt 13620
ttgccgttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680
gccactggag cacctcaaaa acaccatcat acactaaatc agtaagttgg cagcatcacc 13740
cataattgtg gtttcaaaat cggctccgtc gatactatgt tatabgcaa ctttgaaaac 13800
aactttgaaa aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg 13860
gagttcgtct tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa 13920
ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980
accgctgcgt aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag 14040
aaaatgaaaa cctatattta aaaatgacgg acagccggta taaagggacc acctatgatg 14100
tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtc 14160
tgcactttga acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc 14220
tttgctcgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280
agtgcacag gctctttcac tccatcgaca tatcggttg tccctatacg aatagcttag 14340
acagccgctt agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg 14400
aaaactggga agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga 14460
cgaaaaagcc cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct 14520
ttgtgaaaga tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcgggaca 14580
agtggatatga cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt 14640
atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt 14700
atattttact ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag 14760

caggagcgca cggacttctt cggcatcaag tgttttggct ctcaggccga gggccacggc 14820

aagtatttgg gcaaggggtc gctggtattc gtgcaggga agattcggaa taccaagtac 14880

gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataagggt ggattatctg 14940

gacaccaagg caccaggcgg gtcaaatcag gaataagggc acattgcccc ggcgtgagtc 15000

ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060

gaactgatcg acgcggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc 15120

atgcgtgctc cccgcgaaac cttccagtcc gtcggctcga tggccagca agctacggcc 15180

aagatcgagc gcgacagcgt gcaactggct cccctgccc tgcccgccc atcgccgcc 15240

gtggagcgtt cgcgtcgtct cgaacaggag gcggcagggt tggcgaagtc gatgaccatc 15300

gacacgcgag gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa 15360

caggtcagcg aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa 15420

atgcagcttt ccttgttcga tattgcgcgc tggccggaca cgatgcgagc gatgccaaac 15480

gacacggccc gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg 15540

caaaacaagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgctcag 15600

ctgcgggccc acgatgacga actggtgtgg cagcagggtg tggagtacgc gaagcgcacc 15660

cctatcgggg agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggtcg 15720

atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780

atgggcttca cgtccgaccg cgttggggcac ctggaatcgg tgtcgctgct gcaccgcttc 15840

cgcgtcctgg accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc 15900

gtcgtgctgt ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg 15960

tcgccgacgg ccgacggat gttcgactat ttcagctcgc accgggagcc gtacccgctc 16020

aagctggaaa ccttccgcct catgtgcgga tccgattcca cccgcgtgaa gaagtggcgc 16080
gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg gcctgggtgga acacgcctgg 16140
gtcaatgatg acctgggtgca ttgcaaacgc tagggccttg tgggggtcagt tccggctggg 16200
ggttcagcag ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact 16260
tgcttcgctc agtatcgctc gggacgcacg gcgcgtctc cgaactgccg ataaacagag 16320
gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 16380
aggatttccg cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg 16440
tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500
tggcattcgg cgcctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg 16560
acggcccaaa ggacgctcac aaggcgcac tgtccggcgt tttcgtggag cccgaacagc 16620
gaggccgagg ggtcgccggg atgctgctgc gggcggtgcc ggcgggttta ttgctcgtga 16680
tgatcgccg acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac 16740
ttaatatctt gctattctgg agcttgttgt ttatttcggg ctaccgcctg ccgggcgggg 16800
tcgccccgac ggtaggcgct gtgcagccgc tgatggctgt gttcatctct gccgctctgc 16860
taggtagccc gatacgattg atggcggtcc tgggggctat ttgcggaact gcgggcgtgg 16920
cgctgttggg gttgacacca aacgcagcgc tagatcctgt cggcgctcga gcgggcctgg 16980
cgggggcggg ttccatggcg ttcggaacog tgctgacctg caagtggcaa cctcccgtgc 17040
ctctgctcac ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag 17100
ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggctggcgt 17160
ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220
tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280

cggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcgggtgag 17340
caatggatag gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc 17400
ttcctcagcg gctttatcca gcgatttcct attatgtcgg catagtcttc aagatcgaca 17460
gcctgtcacg gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg 17520
agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct 17580
ccgcgagatc atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc 17640
ggtaacatga gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggaact 17700
gatgggctgc ctgtatcgag tgggtathtt gtgccgagct gccggtcggg gagctgttgg 17760
ctggctggtg gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac 17820
acattgcgga cgtttttaat gtactggggt ggtttttctt ttcaccagtg agacgggcaa 17880
cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggc ccacgctggt 17940
ttgccccagc aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat 18000
aaatcaaaag aatagcccg gataggggtg agtggtgttc cagtttgga caagagtcca 18060
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120
ccactacgtg aaccatcacc caaatcaagt tttttggggt cgaggtgccg taaagcacta 18180
aatcggaacc ctaaaggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg 18240
gcgagaaagg aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg 18300
ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 18420
ggccagtga ttcgagctcg gtaccggg 18449

<210> 50
<211> 18617
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 50
ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgcg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcgcca ccgtaacct gtcttttaac ctgcttttaa accaatatatt ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaac cctcccgccc cgctaacgcg ggctctccat cccccaggg gctgcgcccc 660

tgggccgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgga cgcccgaag cattgacgtg 780

ccgcagggtg tggcatcgac attcagcgac cagggtgccg gcagtgaggg cggcggcctg 840

ggtggcgccc tgcccttcac ttgggccgtc ggggcattca cggacttcat ggcggggccc 900

gcaattttta ctttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatata gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatacttata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatcac cagtcaaag 1560

ggtgacagca ggctcataag acgcccagc gtgcgcatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttcggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgctccatt tccgcgcaga cgatgacgtc actgcccggc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860
tgatthttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100
aaatcggtc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggttggaag gaaagctgcc tgttccaaag gtcttgcaact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactocattt aaagatccgc gcgagctgta tgatthttta aagacggaaa agcccgaaga 2820
ggaacttgte ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatthtt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttcgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccgccgtg agtcggggca atcccgcaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgctg gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatggtcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgcg 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380
gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860
catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccga cagcgaggcc gaggggtcgc 4980
cggatatgtg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata tttcgctatt 5100
ctggagcttg ttgtttattt cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtcctggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgtgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgcttgcaa ctggcggcgg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactogaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctcgg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgga gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgcgcct tacaacggct ctcccgctga cgccgtcccg gactgatggg ctgcctgtat 6060
cgagtgggta ttttgtccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg gggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300
aaatcctggt tgatggtggt tccgaaatcg gcaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggccacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcggtaacc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900
gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320
tcctttgggt ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaacctt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcg catgaggttt 7440
tcgaaataca tccggatgtc gaaggcttgg ggcacctgag ttggttgaat ttagaacgtg 7500
gcactattga tcatccgata gctctgcaa gggcggttga caatgcaagt caaacgttgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaata 7920
tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980
cctagctgat tctggagtga cccagagggg catgacttga gcctaaaata cggcgctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgctg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct totccgaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640
tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggtcgttgc gtcagtccaa catttggtgc catattttcc tgctctccc accagctgct 8760
cttttctttt ctctttcttt toccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880
tttgaacctt tcagttcgag ctttcccact tcatcgagc ttgactaaca gctacccgc 8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000
aaagtctgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgtcttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120
ctacaaagat cgttatgttt atcggcaactt tgcacggcc gcgctccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc gtgcacaggg 9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300
ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtocaa gcagcaaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct cggcggaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctott ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcggtgact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800

tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860

atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920

aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt tttttttat 10980

acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040

atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100

tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat tttactaatt 11160

cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220

ttgattacac tctaaatctc gatgaccoga ccaaaaagag cgaacaaaga aataatcttg 11280

tgcatcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340

tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400

tatgctgtcg aagctgcagt caatcagcgt caaggccgc gcggttgaa tagccgcga 11460

catcacgagg cccaaagtct gcctgcatgc tcagcgggct tcgttagttc ggctgcgagt 11520

ggcagcacca cagacagagg aggcgctggg aaccgtgcag gctgccggcg cgggcgatga 11580

gcacagcgcc gatgtagcac tccagcagct tgaccgggct atcgagagc gtcgtgccg 11640

gcgcaaacgg gagcagctgt cataccaggc tgccgccatt gcagcatcaa ttggcgtgtc 11700

aggcattgcc atcttcgcca cctacctgag atttgccatg cacatgaccg tgggcggcgc 11760

agtgccatgg ggtgaagtgg ctggcactct cctcttggtg gttggtggcg cgctcggcat 11820

ggagatgtat gcccgctatg cacacaaagc catctggcat gagtcgcctc tgggctggct 11880

gctgcacaag agccaccaca cacctcgac tggacccttt gaagccaacg acttgtttgc 11940
aatcatcaat ggactgcccg ccatgctcct gtgtaccttt ggcttctggc tgcccaacgt 12000
cctgggggcg gcctgctttg gagcggggct gggcatcacg ctatacggca tggcatatat 12060
gtttgtacac gatggcctgg tgcacaggcg ctttcccacc gggcccatcg ctggcctgcc 12120
ctacatgaag cgctgacag tggcccacca gctacaccac agcggcaagt acggtggcgc 12180
gccctggggg atgttcttgg gtccacagga gctgcagcac attccagggt cggcggagga 12240
ggtggagcga ctggctcctgg aactggactg gtccaagcgg tagaagcttg agattaaaat 12300
agataaggaa aagaaagtga aaagaaattc ggaagcatgg cacattcttc tttttataaa 12360
tacatgcctg actttctttt tccatcgata tgatatatgc atatgataga tatacaagca 12420
atcttcttca aggagtttga aattttgtcc tccaggagca aaaaaagtt ttttttata 12480
catgtttgta cacaagaata gttaccaatt tgctttggtc ttacgtgctg caagtttata 12540
tcgttttcaa tttctttgtc tttacatttt ctttgtcctt tatctttcct catttagtct 12600
ttgggagaat taggaaaagg gagcggaag gtaagaaatg cttgcgtatt ttactaatc 12660
ggcaaacatc caatttggca aacagcagcc tgtgcaacgc tctcgagatg acagtatctt 12720
tgattacact ctaaattctg atgacccgac caaaaagagc gaacaaagaa ataatcttgt 12780
gcattcgaat atgatggaag attttttccc ccttattcta aatgttgaca tagcgtgtat 12840
gttatataaa caaaaagaaa ttgtacaaac tttcttttct tctcttttta ttttatctct 12900
atgatccagt tagaacaacc actcagtcac caagcaaac tgactccagt actgagaagt 12960
aaatctcagt ttaaggggct tttcattgct attgtcattg ttagcgcgtg ggtcattagc 13020
ctgagtttat tactttccct tgacatctca aagctaaaat tttggatgtt attgcctgtt 13080
atactatggc aaacattttt atatacggga ttatttatta catctcatga tgccatgcat 13140

ggcgtagtat ttccccaaaa caccaagatt aatcatttga ttggaacatt gaccctatcc 13200
ctttatggtc ttttaccata tcaaaaacta ttgaaaaaac attgggttaca ccaccacaat 13260
ccagcaagct caatagaccc ggattttcac aatggtaaac accaaagttt ctttgcttgg 13320
tattttcatt ttatgaaagg ttactggagt tggggggcaaa taattgogtt gactattatt 13380
tataactttg ctaaatacat actccatata ccaagtata atctaactta cttttgggtg 13440
ctaccctcgc ttttaagttc attacaatta ttctattttg gtactttttt accccatagt 13500
gaaccaatag ggggttatgt tcagcctcat tgtgccccaa caattagccg tcctatttgg 13560
tggtcattta tcacgtgcta tcattttggc taccacgagg aacatcacga atatcctcat 13620
atttcttggg ggcagttacc agaaatttac aaagcaaat agaagcttgg cgtaatcatg 13680
gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc 13740
cggaagcata aagtgtaaag cctgggggtgc ctaatgagt agctaactca cattaattgc 13800
gttgcgctca ctgccgctt tccagtcggg aaacctgtcg tgccagctgc attaatgaat 13860
cggccaacgc gcggggagag gcggtttgcg tattgggcca aagacaaaag ggcgacattc 13920
aaccgattga gggaggggaag gtaaattatt acggaaatta ttcattaaag gtgaattatc 13980
accgtcaccg acttgagcca tttgggaatt agagccagca aaatcaccag tagcaccatt 14040
accattagca aggccggaaa cgtcaccaat gaaaccatcg atagcagcac cgtaatcagt 14100
agcgacagaa tcaagtttgc ctttagcgtc agactgtagc gcgttttcat cggcattttc 14160
ggcatagcc cccttattag cgtttgccat cttttcataa tcaaaatcac cggaaccaga 14220
gccaccaccg gaaccgcctc cctcagagcc gccaccctca gaaccgccac cctcagagcc 14280
accaccctca gagccgccac cagaaccacc accagagccg ccgccagcat tgacaggagg 14340
cccgatctag taacatagat gacaccgcgc gcgataattt atcctagttt gcgcgctata 14400

ttttgttttc tatcgctat taaatgtata attgctggac tctaatacata aaaacccatc 14460

tcataaataa cgtcatgcat tacatgttaa ttattacatg cttaacgtaa ttcaacagaa 14520

attatatgat aatcatcgca agaccggcaa caggattcaa tcttaagaaa ctttattgcc 14580

aaatgtttga acgatcgggg atcatccggg tctgtggcgg gaactccacg aaaatatccg 14640

aacgcagcaa gatatcgcg tgcattctcg tcttgccctg gcagtcgccc cgcacgccgt 14700

tgatgtggac gccgggccc atcatattgt cgctcaggat cgtggcggtg tgcttgctcg 14760

ccgttgctgt cgtaatgata tcggcacctt cgaccgcctg ttccgcagag atcccgtggg 14820

cgaagaactc cagcatgaga tccccgcgct ggaggatcat ccagccggcg tcccggaaaa 14880

cgattccgaa gcccaacctt tcatagaagg cggcggtgga atcgaaatct cgtgatggca 14940

ggttgggcgt cgcttggtcg gtcatttcga accccagagt cccgctcaga agaactcgtc 15000

aagaaggcga tagaaggcga tgcgctgca atcgggagcg gcgataccgt aaagcacgag 15060

gaagcggcga gccatttcg cgccaagctc ttacgcaata tcacgggtag ccaacgctat 15120

gtcctgatag cggtcgccca caccagccg gccacagtcg atgaatccag aaaagcggcc 15180

attttccacc atgatattcg gcaagcaggc atcgccatgg gtcacgacga gatcatcgcc 15240

gtcgggcatg cgcgccttga gcctggcgaa cagttcggtt ggcgcgagcc cctgatgctc 15300

ttcgtccaga tcatcctgat cgacaagacc ggcttccatc cgagtacgtg ctgctcgat 15360

gcgatgtttc gcttggtggt cgaatgggca ggtagccgga tcaagcgtat gcagccgccg 15420

cattgcatca gccatgatgg atactttctc ggcaggagca aggtgagatg acaggagatc 15480

ctgccccggc acttcgcca atagcagcca gtcccttccc gcttcagtga caacgtcgag 15540

cacagctcg caaggaacgc ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg 15600

cagttcatc agggcacccg acaggtcggt cttgacaaaa agaaccgggc gccctgcgc 15660

tgacagccgg aacacggcgg catcagagca gccgattgtc tgttggtgcc agtcatagcc 15720
gaatagcctc tccaccaag cggccggaga acctgctgc aatccatctt gttcaatcat 15780
gcgaaacgat ccagatccgg tgcagattat ttggattgag agtgaatatg agactctaata 15840
tggataccga ggggaattta tggaacgtca gtggagcatt tttgacaaga aatatttgct 15900
agctgatagt gaccttaggc gacttttgaa cgcgcaataa tggtttctga cgtatgtgct 15960
tagctcatta aactccagaa acccgcggtc gagggtctcc ttcaacgttg cggttctgtc 16020
agttccaaac gtaaacggc ttgtcccgcg tcatcgccgg gggtcataac gtgactccct 16080
taattctccg ctcatgatca gattgtcgtt tccgccttc agtttaaact atcagtgttt 16140
gacaggatat attggcgggt aaacctaaga gaaaagagcg tttattagaa taatcggata 16200
tttaaaagg cgtgaaaagg tttatcgtt cgtccatttg tatgtgcatg ccaaccacag 16260
ggttccccag atctggcgcc ggccagcgag acgagcaaga ttggccgccc cccgaaacga 16320
tccgacagcg cgccagcac aggtgcgcag gcaaattgca ccaacgcata cagcgccagc 16380
agaatgccat agtgggagg gacgtcgtt gagtgaacca gatcgcgag gaggccggc 16440
agcaccggca taatcaggcc gatgccgaca gcgtcgagcg cgacagtgtc cagaattacg 16500
atcaggggta tgttgggtt cacgtctggc ctccggacca gcctccgtg gtccgattga 16560
acgcgcggat tctttatcac tgataagttg gtggacatat tatgtttatc agtgataaag 16620
tgtcaagcat gacaaagttg cagccgaata cagtgatccg tgccgccctg gacctgttga 16680
acgaggtcgg cgtagacggc ctgacgacac gaaaactggc ggaacggttg ggggttcagc 16740
agccggcgct ttactggcac ttcaggaaca agcggcgct gctcgacgca ctggccgaag 16800
ccatgctggc ggagaatcat acgcattcgg tgccgagagc cgacgacgac tggcgctcat 16860
ttctgatcgg gaatccccg agcttcaggc aggcgctgct cgcctaccg gatggcgcg 16920

gcattccatgc cggcacgcga cggggcgcac cgcagatgga aacggccgac gcgcagcttc 16980
gcttcctctg cgaggcgggt ttttcggccg gggacgccgt caatgcgctg atgacaatca 17040
gctacttcac tgttggggcc gtgcttgagg agcaggcccg cgacagcgat gccggcgagc 17100
gcggcggcac cgttgaaacag gctccgctct cgcgctgtt gcgggcccgc atagacgcct 17160
tcgacgaagc cggtcggac gcagcgttcg agcagggact cgcggtgatt gtcgatggat 17220
tggcgaaaag gaggctcgtt gtcaggaacg ttgaaggacc gagaaagggt gacgattgat 17280
caggaccgct gccggagcgc aaccactca ctacagcaga gccatgtaga caacatcccc 17340
tccccctttc caccgcgtca gacgcccgtc gcagcccgt acgggctttt tcatgccctg 17400
ccctagcgtc caagcctcac ggccgcgctc ggccctctctg gcggccttct ggcgctcttc 17460
cgcttcctcg ctactgact cgctgcgctc ggtcgcttcg ctgcggcgag cggtatcagc 17520
tcactcaaag gcggaatac ggttatccac agaatacagg gataacgcag gaaagaacat 17580
gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt 17640
ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg 17700
aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc 17760
tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt 17820
ggcgcttttc cgctgcataa ccctgcttcg gggtcattat agcgattttt tcggtatata 17880
catccttttt cgcacgatat acaggatttt gccaaagggt tcgtgtagac tttccttggt 17940
gtatccaacg gcgtcagccg ggcaggatag gtgaagtagg cccaccgcg agcgggtggt 18000
ccttcttcac tgtcccttat tcgcacctgg cggtgctcaa cgggaatcct gctctgcgag 18060
gctggccggc taccgccggc gtaacagatg agggcaagcg gatggctgat gaaaccaagc 18120
caaccaggaa gggcagccca cctatcaagg tgtactgcct tccagacgaa cgaagagcga 18180

ttgaggaaaa ggccggcggcg gccggcatga gcctgtcggc ctacctgctg gccgtcggcc 18240
agggctacaa aatcacgggc gtcgtggact atgagcacgt ccgcgagctg gcccgcatca 18300
atggcgacct gggccgcctg ggcggcctgc tgaaactctg gctcaccgac gaccgcgca 18360
cggcgcgggtt cgggtgatgcc acgatcctcg ccctgctggc gaagatcgaa gagaagcagg 18420
acgagcttgg caaggatcatg atgggcgtgg tccgcccagag ggcagagcca tgactttttt 18480
agccgctaaa acggccgggg ggtgcgcgtg attgccaagc acgtcccat gcgctccatc 18540
aagaagagcg acttcgcgga gctggtgaag tacatcacccg acgagcaagg caagaccgag 18600
cgccctttgcg acgctca 18617

<210> 51

<211> 18333

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 51

ccgggctggt tgcctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gggccgccg gcgttggtga tacctcgcg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcgcca ccgtaacct gtcttttaac ctgcttttaa accaatattt ataaccttg 540

ttttaacca gggctgcgcc ctgtgcgct gaccgcgcac gccgaagggg ggtgcccccc 600

cttctgaac cctcccgcc cgtaacgcg ggcctcccat cccccaggg gctgcgcccc 660

tcggccgca acggcctcac ccaaaaatg gcagcgtgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcgc cgcgcgca cggccggaag cattgacgtg 780

ccgcagggtc tggcatcgac attcagcgac cagggtccgg gcagtgaggg cggcggcctg 840

ggtggcgcc tgccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatgc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgctcatt tccgcgcaga cgatgacgtc actgcccggc 1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggttca 2100
aaatcggtc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220
aattagcttc ttggggatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccatth aaagatccgc gcgagctgta tgatthttta aagacggaaa agcccgaaga 2820

ggaacttgte ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgtc cggctgatca gggaggatat cggggaagaa cagtatgtcg agctatthtt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgtthtag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgat caagtgttht ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggctcgtggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccgcggtg agtcggggca atcccgaag 3360

gagggatgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggthttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgcctcggc cgcctgggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggte agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatgtcaa acgctagggc cttgtggggt cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgctatt 5100

ctggagcttg ttgtttatctt cggctctaccg cctgcccggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtcctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgcttgcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

tctcttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcacaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctcgg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgca gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgtga cgcgctccg gactgatggg ctgcctgtat 6060

cgagtgggta ttttggtccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cggctccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaacgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcca 6900

gtcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgtga aggcgctttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtgc gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaate 7920

tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcggtga 7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaate cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggtg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtcc 8640

tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggctggtgc gtcagtccaa catttgttgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac tttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgagc ttgactaaca gctaccccg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctctgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacagg 9240

tgtcacgttg caagacctgc ctgaaaccga actgccgct gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt ggcccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggctgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggtccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctctgttact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800
tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860
atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920
aatcttcttc aaggagtgtt aaattttgtc ctccaggagc aaaaaaagt tttttttat 10980
acatgtttgt acacaagaat agttaccaat ttgctttggg cttacgtgct gcaagtttat 11040
atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100
tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgtat ttactaatt 11160
cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220
ttgattacac tctaaatctc gatgaccgga caaaaagag cgaacaaaga aataatcttg 11280
tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340

tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400
tatgttggtg atttggaatg ccctgatcgt tttcgttacc gtgattggca tggaagtgat 11460
tgctgcactg gcacacaaat acatcatgca cggctggggt tggggatggc atctttcaca 11520
tcatgaaccg cgtaaagggt cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580
attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgctcc agtggattgg 11640
cgcaggtatg acggcgtatg gattactcta ttttatgggt cacgacgggc tggtgcatca 11700
acgttggcca ttccgctata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760
ccgtatgcat cagccgtca ggggcaaaga aggttgtgtt tcttttggct tcctctatgc 11820
gccgcccctg tcaaaacttc aggcgacgct ccgggaaaga catggcgcta gagcgggcgc 11880
tgccagagat gcgcagggcg gggaggatga gcccgcatcc gggaagtaag ggcctgacca 11940
gaggcggcca gcagcagcgt taatttttcg ggcgtggctg ttgactgccg ctgatcccaa 12000
agcttgagat taaaatagat aaggaaaaga aagtgaaaag aaattcggaa gcatggcaca 12060
ttcttctttt tataaatata tgcctgactt tctttttcca tcgatatgat atatgcatat 12120
gatagatata caagcaatct tcttcaagga gtttgaaatt ttgtcctoca ggagcaaaaa 12180
aaagtttttt ttatatacatg tttgtacaca agaatagtta ccaatttgct ttggtcttac 12240
gtgctgcaag tttatatcgt tttcaatttc tttgtcttta cattttcttt gtcctttatc 12300
tttctcatt tagtctttgg gagaattagg aaaagggagc ggaaaggtaa gaaatgcttg 12360
cgtattttac taattcggca aacatccaat ttggcaaaca gcagcctgtg caacgctctc 12420
gagatgacag tatctttgat tacactctaa atctcgatga cccgaccaa aagagcgaac 12480
aaagaaataa tcttgtgcat tcgaatatga tggaagattt tttccccctt attctaaatg 12540
ttgacatagc gtgtatgtta tataaacaaa aagaaattgt acaaactttc ttttcttctc 12600

ttttttatatt atctctatga tccagttaga acaaccactc agtcatcaag caaaactgac 12660
tccagtactg agaagtaa atctcagtttaa ggggcttttc attgctattg tcattgttag 12720
cgcatgggtc attagcctga gtttattact ttcccttgac atctcaaagc taaaattttg 12780
gatgttattg cctgttatac tatggcaa acattttatat acgggattat ttattacatc 12840
tcatgatgcc atgcatggcg tagtatttcc ccaaaacacc aagattaatc atttgattgg 12900
aacattgacc ctatcccttt atggctcttt accatatcaa aaactattga aaaaacattg 12960
gttacaccac cacaatccag caagctcaat agaccggat tttcacaatg gtaaacacca 13020
aagtttcttt gcttgggtatt ttcattttat gaaagggtac tggagttggg ggcaaataat 13080
tgcggtgact attatattata actttgctaa atacatactc catatcccaa gtgataatct 13140
aacttacttt tgggtgctac cctcgctttt aagttcatta caattattct attttggtac 13200
ttttttaccc catagtgaac caataggggg ttatgttcag cctcattgtg cccaacaat 13260
tagccgtcct atttgggtgt catttatcac gtgctatcat tttggctacc acgaggaaca 13320
tcacgaatat cctcatattt cttgggtggca gttaccagaa atttacaag caaaatagaa 13380
gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcaaatc 13440
cacacaacat acgagccgga agcataaagt gtaaagcctg ggggtgcctaa tgagttagct 13500
aactcacatt aattgcgttg cgtcactgc ccgctttcca gtcgggaaac ctgtcgtgcc 13560
agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggccaaaga 13620
caaaagggcg acattcaacc gattgagggg ggaaggtaa atattgacgg aaattattca 13680
ttaaaggtga attatcaccg tcaccgactt gagccatttg ggaattagag ccagcaaat 13740
caccagtagc accattacca ttagcaaggc cggaacgct accaatgaaa ccatcgatag 13800
cagcacgta atcagtagcg acagaatcaa gtttgccttt agcgtcagac tgtagcgcgt 13860

tttcatcggc attttcggtc atagccccct tattagcggt tgccatcttt tcataatcaa 13920

aatcaccgga accagagcca ccaccggaac cgctccctc agagccgcca ccctcagaac 13980

cgccaccctc agagccacca ccctcagagc cgccaccaga accaccacca gagccgccc 14040

cagcattgac aggaggcccg atctagtaac atagatgaca ccgcgcgcga taatttatcc 14100

tagtttgccg gctatatttt gttttctatc gcgtattaaa tgtataattg cgggactcta 14160

atcataaaaa cccatctcat aaataacgtc atgcattaca tgttaattat tacatgctta 14220

acgtaattca acagaaatta tatgataatc atcgcaagac cggcaacagg attcaatctt 14280

aagaaacttt attgccaaat gtttgaacga tcggggatca tccgggtctg tggcgggaac 14340

tccacgaaaa tatccgaacg cagcaagata tcgcggtgca tctcggtctt gcctgggcag 14400

tcgccgccga cgccgttgat gtggacgccg ggcccgatca tattgtcgct caggatcgtg 14460

gcgttgtgct tgtcggccgt tgctgtcgta atgatatcgg caccttcgac cgctgttcc 14520

gcagagatcc cgtgggcgaa gaactccagc atgagatccc cgcgctggag gatcatccag 14580

ccggcgctccc ggaaaacgat tccgaagccc aacctttcat agaaggcggc ggtggaatcg 14640

aaatctcgtg atggcaggtt gggcgctcgct tggtcggtca tttcgaacct cagagtcccc 14700

ctcagaagaa ctcgtcaaga aggcgataga aggcgatgcg ctgcgaatcg ggagcggcga 14760

taccgtaaag cagcaggaag cggtcagccc attcgccgcc aagctcttca gcaatatcac 14820

gggtagccaa cgctatgtcc tgatagcggc ccgccacacc cagccggcca cagtcgatga 14880

atccagaaaa gcggccattt tccacatga tattcggcaa gcaggcatcg ccatgggtca 14940

cgacgagatc atcgccgctg ggcgatgcgc ccttgagcct ggcgaacagt tcggctggcg 15000

cgagccctg atgctcttcg tccagatcat cctgatcgac aagaccggct tccatccgag 15060

tacgtgctcg ctcgatgcga tgtttcgctt ggtggtcgaa tgggcaggta gccggatcaa 15120

gcgtatgcag ccgccgcatt gcatcagcca tgatggatac tttctcggca ggagcaaggt 15180
gagatgacag gagatcctgc cccggcactt cgccaatag cagccagtcc cttcccgtt 15240
cagtgacaac gtcgagcaca gctgcgcaag gaacgcccg cgtggccagc cacgatagcc 15300
gcgctgcctc gtccctgcagt tcattcaggg caccggacag gtcggctcttg acaaaaagaa 15360
ccgggcgccc ctgcgctgac agccggaaca cggcggcatc agagcagccg attgtctgtt 15420
gtgcccagtc atagccgaat agcctctcca cccaagcggc cggagaacct gcgtgcaatc 15480
catcttgctc aatcatgcga aacgatccag atccggtgca gattatttgg attgagagtg 15540
aatatgagac tctaattgga taccgagggg aatttatgga acgtcagtgg agcatttttg 15600
acaagaaata tttgctagct gatagtgacc ttagggcact tttgaacgag caataatggt 15660
ttctgacgta tgtgcttagc tcattaaact ccagaaacct cgggctgagt ggctccttca 15720
acgttgccgt tctgtcagtt ccaaagctaa aacggcttgt cccgcgtcat cggcgggggt 15780
cataacgtga ctcccttaat tctccgctca tgatcagatt gtcgtttccc gccttcagtt 15840
taaactatca gtgtttgaca ggatatattg gcgggtaaac ctaagagaaa agagcgttta 15900
ttagaataat cggatattta aaagggcgtg aaaagggtta tccgttcgtc catttgatg 15960
tgcatgcaa ccacaggggt cccagatct ggccggcc agcgagacga gcaagattgg 16020
ccgccgcccg aaacgatccg acagcgcgcc cagcacaggt gcgcaggcaa attgcacaa 16080
cgcatacagc gccagcagaa tgccatagtg ggcggtgacg tcgttcgagt gaaccagatc 16140
gcgcaggagg cccggcagca cgggcataat caggccgatg ccgacagcgt cgagcgcgac 16200
agtgctcaga attacgatca ggggtatggt ggggtttcacg tctggcctcc ggaccagcct 16260
ccgctgggtcc gattgaacgc gcggattctt tatcactgat aagttggtgg acatattatg 16320
tttatcagtg ataaagtgtc aagcatgaca aagttgcagc cgaatacagt gatccgtgcc 16380

gccctggacc tgttgaacga ggtcggcgta gacggtctga cgacacgcaa actggcgga 16440
cggttggggg ttcagcagcc ggcgctttac tggcacttca ggaacaagcg ggcgctgctc 16500
gacgcactgg ccgaagccat gctggcggag aaacatacgc attcggtgcc gagagccgac 16560
gacgactggc gtcattttct gatcgggaat gccgcagct tcaggcaggc gctgctcgcc 16620
taccgcgatg gcgcgcgcac ccatgccggc acgcgaccgg gcgcaccgca gatggaaacg 16680
gccgacgcgc agcttcgctt cctctgcgag gcgggttttt cggccgggga cgccgtcaat 16740
gcgctgatga caatcagcta cttcactgtt ggggccgtgc ttgaggagca ggccggcgac 16800
agcgatgccg gcgagcgcgg cggcaccgtt gaacaggctc cgctctcgcc gctgttgccg 16860
gccgcgatag acgccttcga cgaagccggt ccggacgcag cgcttcgagca gggactcgcg 16920
gtgattgtcg atggattggc gaaaaggagg ctcgttgtca ggaacgttga aggaccgaga 16980
aagggtgacg attgatcagg accgctgccg gagcgcaacc cactcactac agcagagcca 17040
tgtagacaac atccccctcc cctttccacc gcgtcagacg ccgtagcag cccgctacgg 17100
gctttttcat gccctgccct agcgtccaag cctcacggcc gcgctcgcc tctctggcg 17160
ccttctggcg ctcttcgct tctcgtca ctgactcgt gcgctcggtc gttcggctgc 17220
ggcgagcggc atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 17280
acgcaggaaa gaacatgtga gaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 17340
cgttgctggc gtttttccat aggctccgcc .cccctgacga gcatcacaaa aatcgacgt 17400
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcggtt ccccctggaa 17460
gctccctcgt gcgctctcct gttccgacct tgccgcttac cggatacctg tccgcctttc 17520
tcccttcggg aagcgtggcg cttttccgct gcataaccct gcttcggggc cattatagcg 17580
attttttcgg tatatccatc ctttttcgca cgatatacag gattttgcca aagggttcgt 17640

gtagactttc cttggtgtat ccaacggcgt cagccgggca ggataggtga agtaggcca 17700
cccgcgagcg ggtgttcctt cttcactgtc ccttattcgc acctggcggt gctcaacggg 17760
aatcctgctc tgcgaggctg gccggctacc gccggcgtaa cagatgaggg caagcggatg 17820
gctgatgaaa ccaagccaac caggaagggc agcccaccta tcaaggtgta ctgccttcca 17880
gacgaacgaa gagcgattga ggaaaaggcg gcggcgggcg gcatgagcct gtcggcctac 17940
ctgctggccg tcggccaggg ctacaaaatc acgggcgtcg tggactatga gcacgtccgc 18000
gagctggccc gcatcaatgg cgacctgggc cgctgggcg gcctgctgaa actctggctc 18060
accgacgacc cgcgcacggc gcggttcggt gatgccacga tcctcgccct gctggcgaag 18120
atcgaagaga agcaggacga gcttggcaag gtcatgatgg gcgtgggccg cccgagggca 18180
gagccatgac ttttttagcc gctaaaacgg ccgggggggtg cgcggtgattg ccaagcacgt 18240
cccatgctgc tccatcaaga agagcgactt cgcgagctg gtgaagtaca tcaccgacga 18300
gcaaggcaag accgagcgcc ttgcgacgc tca 18333

<210> 52

<211> 17

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3)..(3)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (9)..(9)

<223> n is a, c, g, or t

<400> 52

gcngarggna thtggtta

17

<210> 53

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3)..(3)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (6)..(6)

<223> n is a, c, g, or t

<400> 53

tcngcnagra adatrtrrtg

20

<210> 54

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 54

aagtgacacc gggtacacgc ttgtctt

27

<210> 55

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 55

gcttatcacc atctgttacc tccttgc

27

<210> 56

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 56

agagagggat ccttaaagtc gaatatcggt gc

32

<210> 57

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 57

agagagggat ccatgtctga tcaaaagaag ca

32

<210> 58

<211> 37

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 58

actttattgg atccttaa at gcgaat atcg ttgctgc

37

<210> 59

<211> 38

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 59

gttccaattg gccacatgaa gagtaagaca ggaaacag

38

<210> 60

<211> 38

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 60

cctgtcttac tcttcatgtg gccaattgga accaacac

38

<210> 61

<211> 38

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 61

ctattttaat catatgtctg atcaaaagaa gcatattg

38

<210> 62
<211> 16103
<212> DNA
<213> Artificial

<220>
<223> Primer

<220>
<221> misc_feature
<222> (3471)..(3471)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3679)..(3679)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3770)..(3770)
<223> n is a, c, g, or t

<400> 62
gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60

cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120

cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180

cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct 240

tcgggaagcc agcgactggt gacctccagc gtccttttaa ggctgccaac agctttctca 300

gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360

gtgggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg 420

atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480

tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc 540

aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga 600

gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660

ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720

tccgatagct ctgcaaaggg cgttgcacaa tgcaagtcaa acgttgctag cagttccagg 780

tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840

caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg 900

cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960

tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca 1020

gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140

gttgaccggt gcctggatct tcctatagaa tcattccttat tcgttgacct agctgattct 1200

ggagtgacct agagggtcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260

aaatgtgacg aactcgtgag ctctgtacag tgaccgggtga ctctttcttg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcatcgagga atatggagct 1500

tcatcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560

cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgatc tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860
ggcagctttg ccccgctctgt ccgcccgggtg tgtcggcggg gttgacaagg tcgttgcgtc 1920
agtccaacat ttgttgccat attttctctgc tctccccacc agctgctctt ttcttttctc 1980
tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040
taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100
gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160
accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
gtctccgacc tgatgcagct ctcgaggggc gaagaatctc gtgctttcag cttcgatgta 2280
ggagggcgctg gatatgtcct gcgggtaaat agctgcgcgc atggtttcta caaagatcgt 2340
tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgtc tgacattggg 2400
gaattcagcg agagcctgac ctattgcac tcccgcctg cagagggtgt cacgttgcaa 2460
gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgaggaggc catggatgcg 2520
atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggaacc gcaaggaatc 2580
ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640
tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700
atgctttggg ccgaggactg cccgaagtc cggcacctcg tgcacgcgga ttccggctcc 2760
aacaatgtcc tgacggacaa tggccgcata acagcggta ttgactggag cgaggcgatg 2820
ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
atggagcagc agacgcgcta cttcgagcgg aggcacccgg agcttgacgg atcgccgcgg 2940
ctccggggct atatgtccg cattggtctt gaccaactct atcagagctt ggttgacggc 3000
aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgtccg atccggagcc 3060

gggactgtcg ggcgtacaca aatcgcccg cgaagcgcg cgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cgcccagca ctggtccgag ggcaaaggaa 3180

tagagtagat gccgaccgag ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300

tggttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ctttctagt 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaacct gtaacgcctt ncaggtccg 3480

gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540

tcaacggtcg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600

atatttgtct ctaattgtac ttgacatgc tcctcttctt tactctgata gcttgactat 3660

gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720

ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttctactaa 3780

gatggtatac aatagtaacc atgcatggtt gcctagttaa tgctccgtaa cacccaatac 3840

gccggccgaa acttttttac aactctcta tgagtcgttt acccagaatg cacaggtaca 3900

cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960

ctccacatct ccactcgacc tgcaggcatg caagcttgag tctatgcct caaaaagta 4020

cggtgctgaa ttcagatatc aatcgctgt tgctaaaatt aacactgtcg ataaagacaa 4080

gcgtgtaacc ggtgtcactt tggaaagcg agaagtcatt gaagccgatg cagtcgtatg 4140

taatgcggat cttgtttatg cttatcacca tctgttacct ccttgcaatt ggacaaagaa 4200

gacattagcc tcaaagaaac tcaattcatc atctatttcg ttttattggc ccatgtcaac 4260

aaaggtgcct caattagacg tacacaatat cttcttggt gaagcctaca aggaaagttt 4320

tgatgagatt ttcaacgact tcggtttgcc ctctgaagct tggcgtaatc atggtcatag 4380

ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc 4440

ataaagtgt aagcctgggg tgcctaata gtgagctaac tcacattaat tgcgttgccg 4500

tcaactgccc ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa 4560

cgcgggggga gaggcggttt gcgtattggg ccaaagacaa aagggcgaca ttcaaccgat 4620

tgagggaggg aaggtaaata ttgacggaaa ttattcatta aagggtgaatt atcacctgca 4680

ccgacttgag ccatttgga attagagcca gcaaaatcac cagtagcacc attaccatta 4740

gcaaggccgg aaacgtcacc aatgaaacca tcgatagcag caccgtaatc agtagcgaca 4800

gaatcaagtt tgcctttagc gtcagactgt agcgcgtttt catcggcatt ttcggtcata 4860

gcccccttat tagcgtttgc catcttttca taatcaaaat caccggaacc agagccacca 4920

ccggaaccgc ctccctcaga gccgccaccc tcagaaccgc caccctcaga gccaccaccc 4980

tcagagccgc caccagaacc accaccagag ccgcccag cattgacagg aggcccgatc 5040

tagtaacata gatgacaccg cgcgcgataa tttatcctag tttgcgcgct atattttgtt 5100

ttctatcgcg tattaatgt ataattgcgg gactctaata ataaaaaacc atctcataaa 5160

taacgtcatg cattacatgt taattattac atgcttaacg taattcaaca gaaattatat 5220

gataatcatc gcaagaccgg caacaggatt caatcttaag aaactttatt gccaaatgtt 5280

tgaacgatcg gggatcatcc gggctctgtg cggaactcc acgaaaatat ccgaacgcag 5340

caagatatcg cgggtgcatct cggctcttgc tgggcagtcg ccgccgacgc cggtgatgtg 5400

gacgccgggc ccgatcatat tgcgctcag gatcgtggcg ttgtgcttgt cggccggtgc 5460

tgtcgtaatg atatcggcac cttcgaccgc ctgttccgca gagatcccgt gggcgaagaa 5520

ctccagcatg agatccccgc gctggaggat catccagccg gcgtcccga aaacgattcc 5580

gaagCccaac ctttcataga aggcggcggt ggaatcgaaa tctcgtgatg gcaggttggg 5640

cgtcgcttgg tcggtcattt cgaaccccag agtcccgtc agaagaactc gtcaagaagg 5700

cgatagaagg cgatgcgtg cgaatcgga gggcgatac cgtaaagcac gaggaagcgg 5760

tcagcccatt cgcgcgaag ctcttcagca atatcacggg tagccaacgc tatgtcctga 5820

tagcgggtccg ccacaccag ccggccacag tcgatgaatc cagaaaagcg gccattttcc 5880

accatgatat tcggcaagca ggcacgcca tgggtcacga cgagatcatc gccgtcgggc 5940

atgcgcgcct tgagcctggc gaacagttcg gctggcgca gccctgatg ctcttcgtcc 6000

agatcatcct gatcgacaag accggcttcc atccgagtac gtgctcgctc gatgcgatgt 6060

ttcgcttggg ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca 6120

tcagccatga tggatacttt ctggcagga gcaaggtag atgacaggag atcctgcccc 6180

ggcacttcgc ccaatagcag ccagtccctt ccgcttcag tgacaacgtc gagcacagct 6240

gcgcaaggaa cgcccgtcgt ggccagccac gatagccgcg ctgcctcgtc ctgcagttca 6300

ttcagggcac cggacaggtc ggtcttgaca aaaagaaccg ggcgccctg cgctgacagc 6360

cggaacacgg cggcatcaga gcagccgatt gtctgttgtg ccagtcata gccgaatagc 6420

ctctccacc aagcggccgg agaacctgcg tgcaatccat cttgttcaat catgcgaaac 6480

gatccagatc cgggtgcagat tatttggatt gagagtgaat atgagactct aattggatac 6540

cgaggggaat ttatggaacg tcagtggagc atttttgaca agaaatattt gctagctgat 6600

agtgacctta ggcgactttt gaacgcgcaa taatggtttc tgacgtatgt gcttagctca 6660

ttaaactcca gaaacccgcg gctgagtggc tccttcaacg ttgcggttct gtcagttcca 6720

aacgtaaaac ggcttgctcc gcgtcatcgg cgggggtcat aacgtgactc ccttaattct 6780

ccgtcatga tcagattgtc gtttcccgcc ttcagtttaa actatcagtg tttgacagga 6840

tatattggcg ggtaaaccta agagaaaaga gcgtttatta gaataatcgg atatttataaa 6900

gggcgtgaaa aggtttatcc gttcgtccat ttgtatgtgc atgccaacca cagggttccc 6960

cagatctggc gccggccagc gagacgagca agattggccg ccgcccgaag cgatccgaca 7020

gcgcgcccag cacagggtcg caggcaaatt gcaccaacgc atacagcgcc agcagaatgc 7080

catagtgggc ggtgacgtcg ttcgagtga ccagatcgcg caggaggccc ggcagcaccg 7140

gcataatcag gccgatgccg acagcgtcga gcgcgacagt gtcagaatt acgatcaggg 7200

gtatgttggg ttccacgtct ggctccgga ccagcctccg ctgggtccgat tgaacgcgcg 7260

gattctttat cactgataag ttggtggaca tattatgttt atcagtgata aagtgtcaag 7320

catgacaaag ttgcagccga atacagtgat ccgtgccgcc ctggacctgt tgaacgaggt 7380

cggcgtagac ggtctgacga cagcgaact ggcggaacgg ttgggggttc agcagccggc 7440

gctttactgg cacttcagga acaagcgggc gctgctcgac gactggccg aagccatgct 7500

ggcggagaat catacgcat cgggtgccgag agccgacgac gactggcgct cttttctgat 7560

cgggaatgcc cgcagcttca ggcaggcgct gctgcctac cgcgatggcg cgcgcatcca 7620

tgccggcacg cgaccgggcg caccgcagat ggaaacggcc gacgcgcagc ttcgcttcct 7680

ctgcgaggcg ggtttttcgg ccggggacgc cgtcaatgcg ctgatgaaa tcagctactt 7740

cactgttggg gccgtgcttg aggagcaggc cggcgacagc gatgccggcg agcgcggcgg 7800

caccgttgaa caggctccgc tctcgccgt gttgcccccc gcgatagacg ccttcgacga 7860

agccggtccg gacgcagcgt tcgagcaggg actcgcggtg attgtcgatg gattggcgaa 7920

aaggaggctc gttgtcagga acgttgaagg accgagaaag ggtgacgatt gatcaggacc 7980

gctgccggag cgcaaccac tcaactacagc agagccatgt agacaacatc cctccccct 8040

ttccaccgcg tcagacgccc gtagcagccc gctacgggct ttttcatgcc ctgccctagc 8100

gtccaagcct cacggccgcg ctcgccctct ctggcggcct tctggcgctc ttccgcttcc 8160

tcgctcactg actcgctgcg ctcggtcggt cggtcgggc gagcgggtatc agctcactca 8220

aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca 8280

aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg 8340

ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagagggtg gcgaaacccg 8400

acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtagc ctctcctggt 8460

ccgaccctgc cgcttacggc atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 8520

ttccgctgca taacctgct tcgggggtcat tatagcgatt ttttcggtat atccatcctt 8580

tttcgcacga tatacaggat tttgccaaag ggttcgtgta gactttcctt ggtgtatcca 8640

acggcgctcag ccgggcagga taggtgaagt agggccacc gcgagcgggt gttccttctt 8700

cactgtccct tattcgacc tggcggtgct caacgggaat cctgctctgc gaggctggcc 8760

ggctaccgcc ggcgtaacag atgagggcaa gcgatggct gatgaaacca agccaaccag 8820

gaagggcagc ccacctatca aggtgtactg ccttcagac gaacgaagag cgattgagga 8880

aaaggcggcg gcggccggca tgagcctgtc ggcctacctg ctggccgtcg gccagggcta 8940

caaatcacg ggcgtcgtgg actatgagca cgtccgcgag ctggcccga tcaatggcga 9000

cctgggccgc ctgggcggcc tgctgaaact ctggctcacc gacgaccgc gcacggcgcg 9060

gttcggtgat gccacgatcc tcgccctgct ggcgaagatc gaagagaagc aggacgagct 9120

tggcaaggtc atgatggcg tggccgccc gagggcagag ccatgacttt tttagccgct 9180

aaaacggccg gggggtgcgc gtgattgcca agcacgtccc catgcgctcc atcaagaaga 9240

gcgacttcgc ggagctgggt aagtacatca ccgacgagca aggcaagacc gagcgccttt 9300

gcgacgctca ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca 9360

aacgcgccag aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga 9420

tacctcgccg aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg 9480

gccgactcac ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg 9540

tggagctggc cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttccccacag 9600

atgatgtgga caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact 9660

actgacagat gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg 9720

gcgcacctat tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt 9780

ttccgcccgt ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt 9840

ataaaccttg tttttaacca gggctgcgcc ctgtgcgctg gaccgcgcac gccgaagggg 9900

ggtgcccccc cttctcgaac cctcccggcc cgtaacgcg ggctcccat cccccaggg 9960

gctgcgcccc tcggccgcga acggcctcac ccaaaaatg gcagcgctgg cagtccttgc 10020

cattgccggg atcggggcag taacgggatg ggcgatcagc ccgagcgca cggccggaag 10080

cattgacgtg ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg 10140

cgggcgccctg ggtggcgcc tgcccttcac ttcggccgtc ggggcattca cggacttcat 10200

ggcggggccg gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct 10260

cgtgttcggg ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg 10320

aggtatgaaa acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa 10380

agctaccaag acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac 10440

aatactgata agataatata tcttttatat agaagatatc gccgtatgta aggatttcag 10500

ggggcaaggc ataggcagcg cgcttatcaa tatactata gaatgggcaa agcataaaaa 10560

cttgcattga ctaatgcttg aaaccagga caataacctt atagcttgta aattctatca 10620

taattgggta atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac 10680

tttgtcatgc agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag 10740

gtgctgcctc agattcaggt tatgccgctc aattcgtgc gtatatcgct tgcgtgattac 10800

gtgcagcttt cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac 10860

cacgtcaaag ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc 10920

gaatacgtgc gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgtg 10980

gcgcgattta gcccgcacat agccccactg ttctgtccatt tccgcgcaga cgatgacgtc 11040

actgcccggc tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa 11100

atcgtgttga ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg 11160

gccatatcaa tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt 11220

tgccatgttt tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg 11280

ttacgcacca ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact 11340

ggagcacctc aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat 11400

tgtggtttca aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt 11460

gaaaaagctg ttttctggta ttttaaggtt tagaatgcaa ggaacagtga attggagttc 11520

gtcttgttat aattagcttc ttggggatc tttaaatact gtagaaaaga ggaaggaaat 11580

aataaatggc taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct 11640

gcgtaaaaga tacggaagga atgtctcctg ctaagggtata taagctgggt ggagaaaatg 11700

aaaacctata tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac 11760

gggaaaagga catgatgcta tggctggaag gaaagctgcc tgttcaaag gtcctgcact 11820

ttgaacggca tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct 11880

cggaagagta tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc 11940

tcaggctctt tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc 12000

gcttagccga attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact 12060

gggaagaaga cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa 12120

agcccgaaga ggaacttgtc ttttcccacg ggcacctggg agacagcaac atctttgtga 12180

aagatggcaa agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtgg 12240

atgacattgc cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg 12300

agctatTTTT tgacttactg gggatcaagc ctgattggga gaaataaaa tattatattt 12360

tactggatga attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag 12420

cgcaccgact tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat 12480

ttgggcaagg ggtcgtggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag 12540

gacggccaga cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc 12600

aaggcaccag gcgggtcaaa tcaggaataa gggcacattg ccccggtg agtcggggca 12660

atcccgcaag gaggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg 12720

atcgacggg ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgct 12780

gcgccccg aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc 12840

gagcgcgaca gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag 12900

cgttcgcgtc gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg 12960

cgaggaacta tgacgaccaa gaagcgaaaa accgcccggc aggacctggc aaaacaggtc 13020

agcgaggcca agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag 13080

ctttccttgt tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg 13140

gcccgcctctg ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac 13200

aaggtcattt tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg 13260

gccgacgatg acgaactggg gtggcagcag gtgttgaggt acgcgaagcg caccctatc 13320

ggcgagccga tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat 13380

ggccggtatt acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc 13440

ttcacgtccg accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc 13500

ctggaccgtg gcaagaaaac gtcccggtgc caggctcctga tcgacgagga aatcgctgtg 13560

ctgtttgctg gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgcgg 13620

acggcccgac ggatgttcga ctatttcagc tcgcaccggg agccgtaccg gctcaagctg 13680

gaaaccttcc gcctcatgtg cggatcggat tccaccgcgc tgaagaagtg gcgcgagcag 13740

gtcggcgaag cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat 13800

gatgacctgg tgcattgcaa acgctagggc cttgtggggg cagttccggc tgggggttca 13860

gcagccagcg ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc 13920

gctcagtatc gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa 13980

aattgacaat tgtgattaag gctcagattc gacggccttg agcggccgac gtgcaggatt 14040

tccgcgagat ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg 14100

agcacgagga gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat 14160

tcggcgcccta catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc 14220

ccaaggacgc tcacaaggcg catctgtccg gcgttttcgt ggagcccga cagcgaggcc 14280

gaggggtcgc cggatatgtg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg 14340

tccgacagat tccaacggga atctggtgga tgcgcattc catcctcggc gcacttaata 14400

tttcgctatt ctggagcttg ttgtttatth cggtctaccg cctgccgggc ggggtcgcg 14460

cgacggtagg cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgt ctgctaggta 14520

gcccgatacg attgatggcg gtccctggggg ctatttgccg aactgccggc gtggcgctgt 14580

tgggtgtgac accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg 14640

cggtttccat ggcgttcgga accgtgctga cccgcaagt gcaacctccc gtgcctctgc 14700

tcacctttac cgcctggcaa ctggcggccg gaggacttct gtcggttcca gtagctttag 14760

tgtttgatcc gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg 14820

gcctgatcgg agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac 14880

ctacagttgt ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga 14940

tgcacaggc cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg 15000

ataggggagt tgatatcgtc aacgttcaact tctaaagaaa tagcgccact cagcttcctc 15060

agcggcttta tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt 15120

cacggttaag cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga 15180

tatttgatca caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga 15240

gatcatccgt gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac 15300

atgagcaaag tctgccgct tacaacggct ctcccgtga cgcgtcccg gactgatggg 15360

ctgcctgtat cgagtgggtga ttttgtgccg agctgccggg cggggagctg ttggctggct 15420

ggtggcagga tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg 15480

cggacgtttt taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg 15540

attgcccttc accgcctggc cctgagagag ttgcagcaag cgggccacgc tggtttgccc 15600

cagcaggcga aaatcctgtt tgatgggtgg tccgaaatcg gcaaaatccc ttataaatca 15660

aaagaatagc ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta 15720
aagaacgtgg actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta 15780
cgtgaaccat cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg 15840
aaccctaaag ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga 15900
aaggaaggga agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg 15960
gcgatcggtg cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag 16020
gcgattaagt tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag 16080
tgaattcgag ctoggtaccc ggg 16103

<210> 63

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 63

ggcgtacttg aaggaaccct taccg 25

<210> 64

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 64

attgatgctc ccggtcaccg tgatt 25

<210> 65
<211> 500
<212> DNA
<213> *Blakeslea trispora*

<400> 65
aatctataca atgctccata gactcacatt gatattgtcg aagatttcga tgctgactta 60

gtagagcaac tacaaaagtt agcagagaag catgatttct taatctttga agaccgcaag 120

tttgagata tcggtatgtg aattctatct attttttttc tgatgtgtgc atggatgact 180

catgatcata ttcttaggta atactgtcaa gcatcaatat ggcaagggcg ttacaagat 240

tgcttcttgg tctcatatta ctaatgctca cacagttcct ggagaaggta ttatcaaggg 300

acttgccgaa gtcggcctcc ctcttggtcg tggcttgctt ttgctagcag aaatgtcatc 360

tcaaggtgca ttaactaagg gtatttacac tgccgaatct gtcaatatgg ctgcccga 420

caaagatttc gtttttggct ttattgcaca acacaaaatg aatcagtatg atgatgagga 480

ttttgttgc atgtcgctg 500

<210> 66
<211> 611
<212> DNA
<213> *Blakeslea trispora*

<400> 66
gagattaaaa tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt 60

ctttttataa atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag 120

atatacaagc aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt 180

ttttttttat acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct 240

gcaagtttat atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc 300

tcatttagtc tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat 360

tttactaatt cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat 420
gacagtatct ttgattacac tctaaatctc gatgaccgga ccaaaaagag cgaacaaaga 480
aataatcttg tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac 540
atagcgtgta tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt 600
attttatctc t 611

<210> 67

<211> 720

<212> DNA

<213> Blakeslea trispora

<400> 67

atgtcaatac tcacttatct ggaatttcac ctctactata cactacctgt ccttgcgcca 60
ttgtgttggc tgctaaagcc gtttcactca cagcaagaca atctcaagta taaattttta 120
atgttgatgg ccgcctctac cgcacgatt tgggacaatt atatcgttta tcatcgcgct 180
tggtgggtact gtccacttg tgttggtgct gtcattggct atgtacctct agaagaatac 240
atgttcttta tcatcatgac tttaatgact gtcgcgttct caaactttgt tatgcgttgg 300
cacttgcata ctttctttat tagaccaac acttcttggga agcaaact attagtaggc 360
cttgtgcctg ttccagcttt attggcaatc acttatcatg cttggcactt gacactgcc 420
aataaacctt cattttatgg ttcatgcac ctttggatatg cttgtcctgt gttggctatt 480
ctttggctgg gtgctggcga atatatcttg cgtcgacctg tggctgtcct tttgtctatt 540
gttatcccta gtgtatacct atgttgggct gatatcgctg ctattagtgc tggcacatgg 600
catatttctc ttagaacaag cactggcaaa atggtagtac ccgatttacc tgtagaagaa 660
tgctgtttt ttactttgat caacacagtc ttggtttttg ctacctgtgc tatagaccgc 720

<210> 68
<211> 1089
<212> DNA
<213> *Blakeslea trispora*

<400> 68
ctgtacaaat catctgttca aaatcaaaac cctaaacaag ccatttcct tttccagcat 60

gtcaaagagc tagcatgggc cttctgtctt cctgaccaa tgctcaaca tgaattgttt 120

gatgatctta ctatcagctg ggatatttta cgtaaagcct caaagtcatt ctatactgca 180

tctgccgttt ttccaagtta tgtacgtcaa gacttgggtg ttctctatgc tttctgcaga 240

gctaccgatg acctgtgcga tgatgaatcc aaatctgttc aagaaagaag agaccaatta 300

gatcttactc gacaatttgt tcgtgatctc tttagccaaa agaccagtgc gcctattgtg 360

attgattggg aattgtatca aaaccaactt cctgcttctt gtatatcagc ctttagagcc 420

tttactcgcc ttgccatgt ccttgaagta gaccctgtag aagaactatt agatgggtac 480

aaatgggatc ttgagcgtcg tcctatcctt gatgaacaag acttggaggc atactctgct 540

tgtgtggcca gtagtgtggg tgaaatgtgc acacgtgtga ttcttgctca agacaaaag 600

gaaaatgatg cttggataat tgaccgtgca cgtgagatgg ggctgggtgct acaatacgtt 660

aacattgctc gagacattgt gactgatagc gagactctgg gtcgatgtta tctgcctcaa 720

caatggctta gaaaagaaga aacagaacaa atacagcaag gcaacgccc tagcctaggt 780

gatcaaagac tgttgggctt gtctctgaag cttgtaggaa aggcagacgc tatcatgggtg 840

agagctaaga agggcattga caagttgccg gcaaactgtc aaggcgggtg acgagctgct 900

tgccaagtat atgctgcaat tggatctgta ctcaagcagc agaagacaac atacctaca 960

agagctcatc taaaaggaag cgaacgtgcc aagattgctc tgttgagtgt atacaacctc 1020

tatcaatctg aagacaagcc tgtggctctc cgtcaagcta gaaagattaa gagttttttt 1080

gttgattag

1089

<210> 69

<211> 611

<212> DNA

<213> *Blakeslea trispora*

<400> 69

agagataaaa taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac 60

atacacgcta tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc 120

acaagattat ttctttgttc gctctttttg gtcgggtcat cgagatttag agtgtaatca 180

aagatactgt catctcgaga gcgttgacac ggctgctgtt tgccaaattg gatgtttgcc 240

gaattagtaa aatacgcaag ctttctttac ctttcgctc ctttttcta attctcccaa 300

agactaaatg aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga 360

tataaacttg cagcacgtaa gaccaaagca aattggtaac tattcttgtg tacaacatg 420

tataaaaaaa aacttttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat 480

tgcttgata tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta 540

tttataaaaa gaagaatgtg ccatgcttcc gaatttcttt tcactttctt ttccttatct 600

attttaatct c 611

<210> 70

<211> 882

<212> DNA

<213> *Haematococcus pluvialis*

<400> 70

atgctgtcga agctgcagtc aatcagcgtc aaggcccgcc gcgttgaact agcccgcgac 60

atcacgcggc ccaaagtctg cctgcatgct cagcgggtgct cgttagtctg gctgcgagtg 120

gcagcaccac agacagagga ggcgctggga accgtgcagg ctgccggcgc gggcgatgag 180
cacagcgccg atgtagcact ccagcagctt gaccgggcta tcgcagagcg tcgtgcccgg 240
cgcaaacggg agcagctgtc ataccaggct gccgccattg cagcatcaat tggcgtgtca 300
ggcattgcca tcttcgccac ctacctgaga ttgccatgc acatgaccgt gggcggcgca 360
gtgccatggg gtgaagtggc tggcactctc ctcttgggtg ttggtggcgc gctcggcatg 420
gagatgtatg cccgctatgc acacaaagcc atctggcatg agtcgcctct gggctggctg 480
ctgcacaaga gccaccacac acctcgcact ggaccctttg aagccaacga cttgtttgca 540
atcatcaatg gactgcccgc catgctcctg tgtacctttg gcttctggct gcccaacgtc 600
ctggggggcg cctgcttttg agcggggctg ggcattcacgc tatacggcat ggcatatatg 660
tttgtacacg atggcctggg gcacaggcgc tttcccacg ggcccatcgc tggcctgccc 720
tacatgaagc gcctgacagt ggcccaccag ctacaccaca gcggcaagta cggtggcgcg 780
ccctggggta tgttcttggg tccacaggag ctgcagcaca ttccagggtc ggcggaggag 840
gtggagcgac tggctcctgga actggactgg tccaagcggg ag 882

<210> 71

<211> 528

<212> DNA

<213> Erwinia uredovora

<400> 71

atgttggtgga tttggaatgc cctgatcggt ttcgttaccg tgattggcat ggaagtgatt 60
gctgcactgg cacacaaata catcatgcac ggctgggggtt ggggatggca tctttcacat 120
catgaaccgc gtaaagggtgc gtttgaagtt aacgatcttt atgccgtggt ttttgctgca 180
ttatcgatcc tgctgattta tctgggcagt acaggaatgt ggccgctcca gtggattggc 240

343/357

gcaggatatga cggcgtatgg attactctat tttatggtgc acgacgggct ggtgcatcaa 300

cgttggccat tccgctatat tccacgcaag ggctacctca aacggttgta tatggcgcac 360

cgtatgcatac acgccgtcag gggcaaagaa ggttgtgttt cttttggctt cctctatgcg 420

ccgcccctgt caaaacttca ggcgacgctc cgggaaagac atggcgctag agcgggcgct 480

gccagagatg cgcagggcgg ggaggatgag cccgcatccg ggaagtaa 528

<210> 72

<211> 762

<212> DNA

<213> Nostoc sp. PCC73102

<400> 72

atgatccagt tagaacaacc actcagtcac caagcaaac tgactccagt actgagaagt 60

aaatctcagt ttaaggggct tttcattgct attgtcattg ttagcgcatg ggtcattagc 120

ctgagtttat tactttccct tgacatctca aagctaaaat tttggatggt attgcctggt 180

atactatggc aaacattttt atatacggga ttatttatta catctcatga tgccatgcat 240

ggcgtagtat ttccccaaa caccaagatt aatcatttga ttggaacatt gaccctatcc 300

ctttatggtc ttttaccata tcaaaaacta ttgaaaaaac attggttaca ccaccacaat 360

ccagcaagct caatagaccc ggattttcac aatggtaaac accaaagttt ctttgcttgg 420

tattttcatt ttatgaaagg ttactggagt tgggggcaaa taattgcgtt gactattatt 480

tataactttg ctaaatacat actccatata ccaagtgata atctaactta cttttgggtg 540

ctaccctcgc ttttaagttc attacaatta ttctattttg gtactttttt accccatagt 600

gaaccaatag ggggttatgt tcagcctcat tgtgcccaaa caattagccg tcctatttgg 660

tggtcattta tcacgtgcta tcattttggc taccacgagg aacatcacga atatcctcat 720

atttcttgggt ggcagttacc agaaatttac aaagcaaat ga 762

<210> 73
<211> 617
<212> DNA
<213> *Haematococcus pluvialis*

<400> 73
taggggtgagg aaccaggcac gctgggtttca cacctcatgc ctgtgataag gtgtggctag 60
agcgatgcgt gtgagacggg tatgtcacgg tcgactggtc tgatggccaa tggcatcggc 120
catgtctggc catcacgggc tggttgcctg ggtgaagggtg atgcacatca tcatgtgcgg 180
ttggaggggc tggcacagtg tgggctgaac tggagcagtt gtccaggctg gcgttgaatc 240
agtgaggggtt tgtgattggc ggttgtgaag caatgactcc gcccatattc tatttgtggg 300
agctgagatg atggcatgct tgggatgtgc atggatcatg gtagtgcagc aaactatatt 360
cacctagggc tgttggtagg atcagggtgag gccttgacac ttgcatgatg tactcgtcat 420
ggtgtgttgg tgagaggatg gatgtggatg gatgtgtatt ctcagacgta gaccttgact 480
ggaggcttga tcgagagagt gggccgtatt ctttgagagg ggaggctcgt gccagaaatg 540
gtgagtggat gactgtgacg ctgtacattg caggcagggtg agatgcactg tctcgattgt 600
aaaatacatt cagatgc 617

<210> 74
<211> 1208
<212> DNA
<213> *Haematococcus pluvialis*

<400> 74
attgtgactg atagcgagac tctgggtcga tggtatctgc ctcaacaatg gcttagaaaa 60
gaagaaacag aacaaatata gcaaggcaac gcccgtagcc taggtgatca aagactgttg 120
ggcttgtctc tgaagcttgt aggaaggca gacgctatca tggtagagagc taagaagggc 180

```

attgacaagt tgccggcaaa ctgtcaaggc ggtgtacgag ctgcttgcca agtatatgct 240
gcaattggat ctgtactcaa gcagcagaag acaacatatc ctacaagagc tcatctaaaa 300
ggaagcgaac gtgccaagat tgctctgttg agtgtataca acctctatca atctgaagac 360
aagcctgtgg ctctccgtca agctagaaag attaagagtt tttttgttga ttagtgaatt 420
tttgttttat ttatgtctga tagttcaata aagagacaac acatacaata taaaatcatt 480
gtcttttaaat gttaatttag tagagtgtaa agcctgcatt tttttgtac gcataaacia 540
tgaattcacc ccgcttctgg tttttaaata attatgtcaa actagggaaa attctttttt 600
ttctcttcgt tctttttttg gcttgttgg gagtcacagg cttgtcttca gattgataga 660
ggttgtatac actcaacaga gcaatcttgg cacgttcgct tcctttttaga tgagctcttg 720
taggatatgt tgtcttctgc tgcttgagta cagatccaat tgcagcatat acttggaag 780
cagctcgtac accgccttga cagtttgccg gcaacttgtc aatgcccttc ttagctctca 840
ccatgatagc gtctgccttt cctacaagct tcagagacaa gcccaacagt ctttgatcac 900
ctaggctacg ggcgttgccct tgctgtatct gttctgtttc ttcttttcta agccattgtt 960
gaggcagata acatcgaccc aacatcctcg agccatacta cagcataaaa ggatacgttt 1020
tctttaacag aaattttacc ttttgttctc agcacatata aaaaaaaga aatttaagat 1080
gagtaggact tccattctct caaaaatttt attcaatoca taaatgaatt atttttggac 1140
aaaaaagaaa gattatgcct gattttctct attttttttt tttttacaac tccaccaata 1200
ctttctag 1208

```

<210> 75

<211> 6316

<212> DNA

<213> Blakeslea trispora

<220>

<221> misc_feature

<222> (2694)..(2694)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (4263)..(4263)

<223> n is a, c, g, or t

<400> 75

aaggatgaag aatccaactc taataaaaat cttatggata tctttgatcg actcaaaaag	60
gctttcaatg ctattgctat taaaaaaaa gagagagaga gaactatgag caaaaggact	120
ctatgccaaag atggcaaaaa ggcaccagaa acccttagtt tattattgca taatccagtc	180
gagctagtac ttctgtagct caagcttaac cgaggatctt ggaatcaact cgtctcgtca	240
ctcttgccga tgatcctaga aatggatatct atggatgtta tactaacatt gttatctttc	300
aaggcctcga agatgttatt gttgcggtga taaataggct gctatgtact gaagttgctc	360
tgtaaaatga atctagttca ctgcctactc agcaaagggt tgtttctaata gtctttaaag	420
aaagaaaaaa agatacatat agactaccct tcctttcaag actgtaatcg agaatcggcc	480
gatggtttat tacaattaga cgctgggaat aagcaaaagg attcatcttt gtaaataaga	540
gactggtgca tatgaaagca aggatcgtat caaggaatag ttttgatcga gcatcaccag	600
caaatgctgc taatgttggc ttcttctttg cttcctgaga ttgaatggga tgtgcctaga	660
gcattgctat ttttaagtgt atactttaga tttgtgtctt tagatttgtg tcattttatt	720
tagtcaagaa agatccccct ttctctatgt atgctaagaa gaaggagcaa gaagtgtatt	780
tacaagttgg aatgagattg aaatattgta cataataata ataaaaagaa aggtagatca	840
aaaaaatgt tctgcctatt gtaagaaatc gggaccaaca ggtgcttgat aaccagaagt	900

agcttccaat tcaggtagag gctctagga caaatacaca attatgacag gaattttctt 960
gttgacttga acactacaag agaaacgggt cagcacaaaa tccgaaaaaa aaaagaaacg 1020
gaccattcat gtcttaccta tctagctctt tgtcttcaat tgcattccat tgctcaacca 1080
cagatacgct tccaattga gtatattgat gaagtgttc ctgcattttt cgcttgacta 1140
attccactac agtcacagtc ttattaatgt tttgtccttt accagtcagg ataatatgat 1200
ctttttgctt cttctatcaa aaaaataatt cttgttttga ataaaaaaaa caaatattta 1260
aagaaactac tttgatgacg gtacctggaa taactcgaga cacacatcta catatgcgtt 1320
gattttattg tggctaattc gaacctcatt ttctgctggt gggggctggt gactttcagt 1380
tgctgagacg tccttcttgc ttcttttata gtcttcact atgattttta tcaagaaagt 1440
aagtcagtga tgattgttac aagctatata tcttgaaaaa gaacagagag gtattattat 1500
cagatgcaac atggttttct gtatcatttt catttcagtt tctctgttca aaaaaaaaaa 1560
gaacactttc tctttccact cctcaaattt tttctgctaa actcctcgca aaacatgtat 1620
ttgctttaaa ctacaagttg caattgtctg atttagcaat ttcaatatgc cttttgtgaa 1680
tccacccaaa aataaacaag tgcttgagta tacttgggtt cagttcaaaa gaaagcaagc 1740
tttttttttt ctttcttggg aaagaaaaaa aaatattgtt gagccatcct ttaccagcag 1800
tatgcgagct acgacatagc tggcttaaca atgactgcaa gcaatagatc gagcttagtc 1860
tttctattgc ttcyttgttt gatctatgtt cggccttacg ctgacctatc caatactcga 1920
gataggcaac aagatttcga acagtaatga aataaatttc ggataacagt tgtggatgag 1980
gaagagaaag cgacttgaac tcgagaaact ttgttgaaat gaaatccgac cttttacgtg 2040
atcatcatgt attatcctct ttttcttttt tttcgtagtg aattacttac tgattgcgct 2100
caagtcgctt ctttataaag aagaaaaaaa aatattagaa ctttcaaaaa atataactga 2160

aaataaaagt gtggctcgga gagcaaatac cacatccttt gtcttcgctt tggtaacacg 2220

gttaataagc cactataggt gaataatgat catttctgag aataaagcgc ggcttgaagc 2280

ttatatccat atcaggattc atattaggca caactcacia ttgaggttcc agaagtgcc 2340

atTTTTTTTT cctgatagcc tgtccaatta agatcaaaaa ccactgagtt ttctctatat 2400

atTTTTTTTT ttcataatc ttaactcttc ttctctctc tctctctctc tctctttttg 2460

gcttgcaaaa aaaatcttta gtaataccaa agaaagcaaa ctttttcctt ttcttatttc 2520

cttgcttggt ttttaatttt tgatttctct atgctttaa taccatttc tttctttctt 2580

ctgctattac ctatcttttc attcctctcc cccctctctc tcttggtcta taaacatcat 2640

gaagtcctct tttaaaagtt cgcttgacat ttatgctgtt tatatacagc atcntgtgtt 2700

ttccaagtgg ttcatctctg cttttgttct ttcgatttcc ctcaacactt atctactgaa 2760

cgcttcgaag caacagccca aagtataat caaaaagggt attgagcggg tagaagtacc 2820

aagtagagaa caacctaat cagtcataaa gccctcctcc aagaaacact cttctcatca 2880

tcagtctgat gtcattcgcc ctcttgatga agtattgggt ttgctcggaa cacccgaggc 2940

cttgactgat gaagagatca tctctattgt tcaagctgggt aaaatggccc cctatgctct 3000

tgaaaaggtc ttgggcgatt tagagcgcgc tgtccatata cgctcgtgctt tgatctcccg 3060

tgactctcgt acgaaaactt tggaagacag tatgcttccc gtgaaaaact atcattatga 3120

taaagtcatg ggtgcttggt gtgaaaatgt cattgggtat atgcctattc cagtaggtgt 3180

cgcaggtaag aagttcaaca agtcgcgata ttgacaagt tgctcatcat ttctgaaaca 3240

ggtcctttgg tgattgatgg tgattctatt catattccca tggcaactac ggaagggtgt 3300

ttagttgctt ctactgccag aggttgtaaa gcaatcaatg ctgggtggtg tgccaacaca 3360

attgttggtg ctgatgggtat gactcgaggt ccttggtgctg aatttcctac aatcactcgc 3420

gctgctgact gtaaacgatg gattgaacaa gaggggtgaag ctatcgtgac cgaggcattc 3480
aattcaactt ctcgttttgc tcgtgttcgt aaattgaaag ttgctcttgc cggtcgtcta 3540
gtctacatcc gtttctctac cactacaggt gatgcaatgg gcatgaacat gatctccaag 3600
ggttgtgaaa aggctttaag caagattgct gagagatata ctgatatgca gatcatttct 3660
ctttctggta actattgtac tgacaagaaa cctgctgcta tcaactggat tgaaggacgt 3720
ggtaaactctg ttgttgctga sgctgtcatc cctgggtacgg ttgtcgaaaa ggtattgaag 3780
acctctgtta gtgctttggg tgagctgaac atctctaaaa acctgggtgg ttctgctatg 3840
gctgggtccg tcgggtggtt taacgctcat gctgctaata ttctaactgc catttacctt 3900
gctactgggc aagatcctgc tcaaaatgta sagagttcta actgtattac tttgatgaaa 3960
gctgtcaatg gcgaaagaga ccttcataatc tcttgtacaa tgccctgtat tgaagtaggc 4020
accattgggtg gtggtactat ttgctctcct caacaagcca tgttggtattt cattgggtgtg 4080
cgtgggtcctc acctaccga acctgggtgcc aatgcccgwc gccttgctcg tgttatctgt 4140
gcctctgtga tggctggtga attgtcttta tgtgcagctt tggctgctgg tcatcttgta 4200
aaggcacaca tggctcataa tcgtaatacc actgctgctg ccgctgttgt tcctgcccct 4260
aanggcatag ttgatgtctc tacacctcct gctacacctg cagaaaagaa tgatcctatt 4320
cctggaagtt gtatcaagtc atagaattaa tattatatat atatcatata caaaaaaag 4380
aaaaaaaaa cactacatct atttatattt ctccatgtac acacacacac acacatataa 4440
aaactcttta tttccaata ttttgctttt ataaataatc ttatttcatt ctaaataaac 4500
tgtttttttt tattaatcat caaacctgc tgagagctgt gcaatatcat ctatgttttc 4560
atgggtttaac tctgggtatcg gwcgagcctc ctctgtactt gaagtttgta ggcagttttt 4620
atttaaggct gctgggtcgat catgatcatc akcaaacctg acagcatgaa gttttgactg 4680

atgagcaatt tctaaggga cagaatctga actctttcgc ttcctactat tgaccatatt 4740
gtcttttaggt ggaatgagtg aatagcgtct tgtcatatgt aacacagaat caacaatatt 4800
ctggtgatga aactcggcca aacatagcgc ctttctcccc caacaattat aataatcaaa 4860
atgagaatga catgtacggt tttcctcgat gacaatatcc aacgtcttgt cataatcctc 4920
tgtgcgyata ccattcatct tttggaagaa cgcacggtag ctctcacaag ctgtcctcag 4980
agagttccgt gccatgtttc ccaatgctcc tggcaagtcg aaatgaagtt gtcgaatctg 5040
gcatgtatg tctacaatgt cgcctgtttc tttcattaga tcaagcattc gtgtagccca 5100
aatgatgtct atgttatgat tttctttcat tccagtaata actatagttt ctcggaat 5160
cgaatgastg atggagtaaa ttcataaaaa gtgcaagtaa tacatacagt gcttgaagaa 5220
atcttggtga gcacgcctat attatgtaat ataggatcga ttctcgaaac tcgacataac 5280
caccaggctt tagcaagcgt tttatttcat tcatgacaag ctattgttaa ttcygtctta 5340
ataaaacaaa atgaaaaaaa cataaccccc tmaaaactta cttccactc ttgattggaa 5400
aaacaggat agacgtgacg catatgtata taatcaaac actcatcagg atagggtaaa 5460
ccattgagca catcgcatg ggtgaagaaa gtattaggag gcttgatggc tgtaggatat 5520
ataggtgcaa tatcaatacc gtaaaactca gcatttggga attctgtagc catctccaga 5580
atccaagtac ctgtgccaca agcaacatca agcactttag gtaagggat acattgttgt 5640
tcttggtgtt gttgttgaca atcacttgag tctgagtttc gttttgattg ttttaatgac 5700
aataattctt ttacagggtg tgagaaatta cgtcaaata gatacttgta aataaaatgc 5760
taaaaataaa aacaatagaa aaaaaaattg acgtcattt cattactatg gaaataactg 5820
caaaatctta ccacttgtag aagtctatct tgctcaatct catcgtttgg cagaatgtat 5880
ttattgttgt agtattgata tcttctacca ttcattgat aactgtcgt tctaagtctc 5940

tgaggtgaag tacttgtagg tgaaggtgga agtgacgcaa ttttgtcaag cttaacagga 6000
 tcctctcggc tacatgtttt ctgcatatca ggaaaatcct gtttatttga aacatcaaca 6060
 gtagatgtgg tgtgatcttt ttgaaaata tcgatgcctt cctttgaaag ccttttgaaa 6120
 ggctctttta acttttttga gtgagagcta cccatgatag cttatgaaga attaaaaaga 6180
 aaaaagcaaa aaaaattaaa aaaaaaaaaa gtagcaaaaa attctgtcgt aattatacaa 6240
 gccaatcaaa atcgaaattc atgcaaggca tagatgttca cgtggatttg atggttgatc 6300
 cttttttttt gcaaga 6316

<210> 76

<211> 1170

<212> DNA

<213> *Thermus thermophilus*

<400> 76

atgaagcgcc tttccctgag ggaggcctgg ccctacctga aagacctcca gcaagatccc 60
 ctgcgcgtcc tgctggcgtg gggccggggc ccccccggc tcttccttcc cctgccccgc 120
 ttccccctgg cctgatctt tgaccccgag ggggtggagg gggcgctcct cgccgagggg 180
 accaccaagg ccaccttcca gtaccggggc ctctcccgcc tcacggggag gggcctcctc 240
 accgactggg gggaaagctg gaaggaggcg cgcaaggccc tcaaagacct cttcctgccg 300
 aagaacgtcc gcggtaccg ggaggccatg gaggaggagg cccgggcctt cttcggggag 360
 tggcgggggg aggagcggga cctggaccac gagatgctcg ccctctccct ggcctcctc 420
 gggcggggcc tcttcgggaa gccctcttcc ccaagcctcg cggagcacgc ccttaaggcc 480
 ctggaccgga tcatggccca gaccaggagc cccctggccc tcttgacct ggccgccgaa 540
 gcccgcttcc ggaaggaccg gggggccctc taccgcgagg cggaagccct catcgccac 600

ccgcccctct cccaccttcc ccgagagcgc gccctgagcg aggccgtgac cctcctgggtg 660
gcggggccacg agacggtggc gagcgccctc acctgggtcct ttctcctcct ctcccaccgc 720
ccggactggc agaagcgggt ggccgagagc gaggaggcgg ccctcgccgc cttccaggag 780
gccctgaggc tctaccccc cgctggatc ctcaccgga ggctggaaag gccctcctc 840
ctgggagagg accggctccc cccgggcacc acctgggtcc tctcccccta cgtgaccag 900
aggctccact tccccgatgg ggaggccttc cggcccgagc gcttctgga ggaaaggggg 960
acccttcgg ggcgctactt cccctttggc ctggggcaga ggctctgcct ggggcgggac 1020
ttcgccctcc tcgagggcc catcgctctc agggccttct tccgccgtt ccgcctagac 1080
cccctccct tccccgggt cctcgcccag gtcacctga ggcccgaagg cgggcttccc 1140
gcgcggccta gggaggaggt gcgggcgtga 1170

<210> 77

<211> 2981

<212> DNA

<213> *Blakeslea trispora*

<400> 77

tctagaattc attccattcg aaaggatcaa cataaccaat ttaatgacta ctagctaattg 60
gatacaaata tacgcacaaa aaaagaaaga attctatgat caaagagaac acagacacag 120
agtgatacat ttaaattggtt aagttcttat gatgttaaaa tggtaaacttt attattgaat 180
taaattgcgaa tatcgttgct gctttgtact tggaaaacgt taggtaaaag ttggttaatg 240
aaagaagcag gagttgtagt atcatctctt gggaagaaat agaaaaagag gaaagtaaca 300
aagtaacaag caagacaata atagatccaa tggttttcgg tcttacgagt ttgttcagga 360
gcatacttct tttggctatc ttgtaacttt cttggtaagg gattctggcc aaagctttta 420
cagacttggt cggaagtaag cttacttcca gcaagaacga taggaacacc agtacctgga 480

tgtgtactac aaagaaaaga gaaatgagta cgtgcgttat taaaaaaaag aaaaaaagag 540

ggcaaaagta ttacctagct ccgacaaaga aaagattatc ataacggttt gtggaatcct 600

tgggtactagg tctgaaccag agaacttgga acacatcatg agaaagacca agaatagaac 660

ctctccaaag gttaaacttg ctttgccaaa cactaggatc attcacttct tcatgttcaa 720

tcaaattagc aaagttgttt actcccaaac gacgttcgat aacttccaga accatcttgc 780

gtgcacgggt taccaactca ggataathtt cttcagcact gtttcctgtc ttactcttca 840

tatggccaat tggaaccaac acaataatgg agtccttggt gggagggtgcg gcagattcat 900

caattcgaga tggaacgttg acatagaatg aagcttcaga gggcaaaccg aagtcgttga 960

aaatctcatc aaaactttcc ttgtaggctt cagccaagaa gatatttgtt acgtctaatt 1020

gaggcacctt tgttgacatg gaccaataaa acgaaataga tgatgaagtg agtttctttg 1080

aggctaattg cttctttgtc caattgcaag gaggtaacag atgggtgataa gcataaacia 1140

gatccgcatt acatacgact gcatcggctt caatgacttc tccgctttcc aaagtgcac 1200

cggttacacg cttgtcttta tcgacagtgt taatttttagc aacaggcgat tgatatctga 1260

attcagcacc gtactttttg gaggcgatag actcaagctt ctgaacaacc atgttgaaac 1320

caccacgagg ataccagata ctttcagcaa actcgggtga ttgtaacaaa ctgtaaactg 1380

ctggagcatc ataaggcgac atactatatt ccaaaaatag aaaatagaac aatgaatac 1440

aaaattcctt tcaacttccc tttttcacat ttctcttttc ccacccccga ccggtctcac 1500

tcattttttt ttcattccac accacgcgtt gtatgtgtac ttaccccata tacattgttt 1560

gaaaagtaaa agccatacgc attttcttgg tttggaaata tttactggct cggtcataga 1620

tcttacaaa caagtgcag cgaaagattt caggcacata ctgaagacga atcaaattcc 1680

aaatggtttc aaagttgcgc ttgatagcaa taaatgtacc ttgttcataa tggacatgtg 1740

tttccttcat gaaatccaag aatctaccaa atccaagggg accctcaata cggccaatt 1800
cgcccttcat cttgggttaa tcggaagaga gttgtacggc atcacgctcg tcaaaatgaa 1860
ccttatagtt attgtcacag cgaagcaaat ccaatgac accaatacgt tcatccaaat 1920
cagcaaatgc atcttcaaaa agcttaggca tcaaatagag tgagggaccc tgatcaaagc 1980
gatgaccatc gtgatgaatg aatgaacaac ggccaccgga aaagtcgttc ttttcaacaa 2040
cagtaactcg aaaaccttca cgagcaagac gaggcagcagt agcagttccg ccaataccgg 2100
caccaatgac aacaatatgc ttcttttgat cagacatgag attaaaatag ataaggaaaa 2160
gaaagtgaag agaaattcgg aagcatggca cattcttctt tttataaata catgcctgac 2220
tttctttttc catcgatatg atatatgcat atgatagata tacaagcaat cttcttcaag 2280
gagtttgaaa ttttgcctc caggagcaaa aaaaagtttt tttttataca tgtttgatca 2340
caagaatagt taccaatttg ctttgggtct acgtgctgca agtttatatc gttttcaatt 2400
tctttgtctt tacatcttct ttgtccttta tctttctca tttagtcttt gggagaatta 2460
ggaaaaggga gcggaaaggt aagaaatgct tgcgtatctt actaatcgg caaacatcca 2520
atttggcaaa cagcagcctg tgcaacgctc tcgagatgac agtatctttg attacactct 2580
aatctcgtat gacccgacca aaaagagcga acaaagaaat aatcttgtgc attcgaatat 2640
gatggaagat tttttccccc ttattctaaa tgttgacata gcgtgtatgt tatataaaca 2700
aaaagaaatt gtacaaactt tcttttcttc tctttttatt ttatctctat gtcaatactc 2760
acttatctgg aatttcatct ctactataca ctacctgtcc ttgcggcatt gtgttggtg 2820
ctaaagccgt ttcactcaca gcaagacaat ctcaagtata aatttttaat gttgatggcc 2880
gcctctaccg catcgatttg ggacaattat atcgtttatc atcgcgcttg gtggtactgt 2940
cctacttgtg ttgtggctgt cattggctat gtacctctag a 2981

<210> 78

<211> 1749

<212> DNA

<213> *Blakeslea trispora*

<400> 78

atgtctgatac aaaagaagca tattgttgta attggtgccc gtattggcgg aactgctact 60

gctgctcgtc ttgctcgtga aggttttcga gttactgttg ttgaaaagaa cgacttttcc 120

ggtggccgtt gttcattcat tcatcacgat ggtcatcgct ttgatcaggg tccctcactc 180

tatttgatgc ctaagctttt tgaagatgca tttgctgatt tggatgaacg tattggtgat 240

catttggtt tgcttcgctg tgacaataac tataagggtc attttgacga cggatgatgcc 300

gtacaactct cttccgattt aaccaagatg aaggggcgaat tggaccgtat tgagggtccc 360

cttggtttg gtagattctt ggatttcata aaggaaacac atgtccatta tgaacaaggt 420

acatttattg ctatcaagcg caactttgaa accatttggg atttgattcg tcttcagtat 480

gtgcctgaaa tctttcgctt gcacttggtt ggtaagatct atgaccgagc cagtaaatat 540

ttccaaacca agaaaatgcg tatggctttt acttttcaaa caatgtatat gggatgtgcg 600

ccttatgatg ctccagcagt ttacagtttg ttacaataca ccgagtttgc tgaaggatc 660

tggatccctc gtgggtggtt caacatggtt gttcagaagc ttgagtatat cgctccaaa 720

aagtacgggtg ctgaattcag atatcaatcg cctgttgcta aaattaacac tgcgataaa 780

gacaagcgtg taaccggtgt cactttggaa agcggagaag tcattgaagc cgatgcagtc 840

gtatgtaatg cggatcttgt ttatgcttat caccatctgt tacctccttg caattggaca 900

aagaagacat tagcctcaaa gaaactcact tcatcatcta tttcgtttta ttggtccatg 960

tcaacaaagg tgcctcaatt agacgtacac aatatcttct tggctgaagc ctacaaggaa 1020

356/357

agttttgatg agattttcaa cgacttcggt ttgccctctg aagcttcatt ctatgtcaac 1080
gttccatctc gaattgatga atctgccgca cctcccaaca aggactecat tattgtgttg 1140
gttccaattg gccatatgaa gagtaagaca ggaacacagt ctgaagaaaa ttatcctgag 1200
ttggtaaacc gtgcacgcaa gatggttctg gaagttatcg aacgtcgttt gggagtaaag 1260
aactttgcta atttgattga acatgaagaa gtgaatgac ctagtgtttg gcaaagcaag 1320
tttaaccttt ggagagggtc tattcttggg ctttctcatg atgtgttcca agttctctgg 1380
ttcagacctc gtaccaagga ttccacaaac cgttatgata atcttttctt tgcggagct 1440
agtacacatc caggtactgg tgctcctatc gttcttgctg gaagtaagct tacttccgac 1500
caagtctgta aaagctttgg ccagaatccc ttaccaagaa agttacaaga tagccaaaag 1560
aagtatgctc ctgaacaaac tcgtaagacc gaaagccatt ggatctatta ttgtcttgct 1620
tggtactttg ttactttcct ctttttctat ttcttcccaa gagatgatac tacaactcct 1680
gcttctttca ttaaccaact ttacctaac gttttccaag taaaagcag caacgatatt 1740
cgcatttaa 1749

<210> 79

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 79

ccgatggcga cgacggaagg ttgtt 25

<210> 80

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 80

catgttcattgcccattgcatcacct

25